Sampling diverse characters improves phylogenies: Craniodental and postcranial characters of vertebrates often imply different trees

Ross C. P. Mounce[1], Robert Sansom[2] and Matthew A. Wills[1]*

[1]*The Milner Centre for Evolution, Department of Biology and Biochemistry, The University of Bath, The Avenue, Claverton Down, Bath BA2 7AY, UK; E-mail: m.a.wills@bath.ac.uk*

[2]*Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester, M13 9PT*

Words in main text : 6588 (including abstract)
Figures : 5
Tables : 2

***Correspondence:** Matthew A. Wills
Department of Biology and Biochemistry
The University of Bath
The Avenue
Claverton Down
Bath BA2 7AY
United Kingdom

Telephone: 01225 383504
Fax: 01225 386779
E-mail: m.a.wills@bath.ac.uk

Sampling diverse characters improves phylogenies: Craniodental and postcranial characters of vertebrates often imply different trees

Ross C. P. Mounce[1], Robert Sansom[2] and Matthew A. Wills[1]

[1]*The Milner Centre for Evolution, Department of Biology and Biochemistry, The University of Bath, The Avenue, Claverton Down, Bath BA2 7AY, UK; E-mail: m.a.wills@bath.ac.uk*

[2]*Faculty of Life Sciences, University of Manchester, Michael Smith Building, Oxford Road, Manchester, M13 9PT*

Morphological cladograms of vertebrates are often inferred from greater numbers of characters describing the skull and teeth than from postcranial characters. This is either because the skull is believed to yield characters with a stronger phylogenetic signal (i.e., contain less homoplasy), because morphological variation therein is more readily atomized, or because craniodental material is more widely available (particularly in the palaeontological case). An analysis of 85 vertebrate datasets published between 2000 and 2013 confirms that craniodental characters are significantly more numerous than postcranial characters, but finds no evidence that levels of homoplasy differ in the two partitions. However, a new partition test based on tree-to-tree distances (as measured by Robinson Foulds metric) rather than tree length reveals that relationships inferred from the partitions are significantly different about one time in three, much more often than expected. Such differences may reflect divergent selective pressures in different body regions, resulting in different localized patterns of homoplasy. Most systematists attempt to sample characters broadly across body regions, but this is not always possible. We conclude that trees inferred largely from either craniodental or postcranial characters in isolation may differ significantly from those that would result from a more holistic approach. We urge the latter.

Despite the increasing importance of molecular and genomic data over the last two decades, morphology still makes an invaluable contribution to vertebrate phylogenetics, and is the only suitable source of data for palaeontological phylogenies (Asher and Müller, 2012; Hillis and Wiens, 2000). Levels of morphological homoplasy in vertebrate groups are generally lower than those amongst their invertebrate counterparts (Hoyal Cuthill *et al.* 2010), suggesting that the signal quality for vertebrates is relatively high. Morphological systematists usually seek to code as many valid characters from as wide a selection of organs and body regions as possible, and to analyse these simultaneously, typically citing the principle of total evidence (Kluge, 1989). Heterogeneity in the performance of characters is often quantified in retrospect, and can also be utilised formally to increase stability in various *post hoc* weighting schemes (Farris 1969; Goloboff 2014; Goloboff *et al.* 2008a). However, coded matrices are usually analysed holistically rather than in partitions defined *a priori*. This is not least because hidden support (the presence of weak signals within putative partitions that become reciprocally reinforcing, and therefore the dominant signal when all characters are considered) is most readily identified in this way (Gatesy *et al.* 1999; Gatesy and Arctander 2000). It is therefore conceded that different body regions can be subject to different levels and patterns of homoplasy (Gaubert *et al.* 2005), such that anatomical subsets of characters can yield significantly different trees. This is especially true when there is strong functional selection in particular organ systems, leading to convergence (Ji *et al.* 1999; Kivell *et al.* 2013; Tseng *et al.*, 2011).

Few cladistic studies have analysed the performance of morphological characters within partitions explicitly (e.g., Sánchez-Villagra and Williams, 1998; Rae, 1999; Gould, 2001; Song and Bucheli, 2010). While analyses may concentrate upon characters of particular types or from particular organ systems, it is still rare that trees are inferred explicitly from subsets of these data (but see Rae, 1999; Vermeij and Carlson, 2000; O'Leary *et al.*, 2003; Poyato-Ariza, 2003; Diogo, 2004; Young, 2005; Clarke and Middleton 2008; Smith, 2010; Farke *et al.*, 2011). Notable exceptions are taphonomic studies that investigate the effects of omitting volatile, soft-part

101 characters with a low fossilisation potential (Sansom *et al.* 2010, 2011; Sansom and Wills, 2013;
102 Pattinson *et al.*, 2014; Sansom 2015).

103 Where levels of homoplasy across morphological character partitions are considered at all,
104 these are usually investigated by comparing distributions of the consistency index (ci: Kluge and
105 Farris, 1969) (Fig. 1). The ci is given simply as the minimum possible number of state changes
106 (states minus characters) divided by the most parsimonious tree length. More appropriate functions
107 underpin most *post hoc* weighting schemes (Goloboff 2014; Goloboff *et al.* 2008a). In the case of
108 insects, Song and Bucheli (2010) found male genital characters to be less homoplastic than
109 characters coding other aspects of form. In brachiopods, Leighton and Maples (2002) revealed that
110 shell characters are significantly more homoplastic than those describing internal anatomy (a
111 disconcerting finding in a group whose fossil record consists largely of hard parts, and in which
112 phylogeny is inferred almost exclusively with recourse to such characters). Similarly, in hedgehogs,
113 Gould (2001) reported significantly higher consistency indices for dental characters compared with
114 those describing other aspects of anatomy. However, she also noted that the optimal trees inferred
115 from the dental characters alone were seriously at odds with those inferred from the other
116 characters. Kangas *et al.* (2004) noted that many morphological characters of mammalian teeth
117 have exceptionally strong developmental interdependence, all being under the correlated control of
118 a small number of genes. Most ambitiously, Sanchez-Villagra and Williams (1998) compared the
119 consistency indices between dental, cranial and postcranial character partitions of eight
120 mammalian datasets, but reported no significant differences. We note that the relative performance
121 of particular characters and character partitions may be contingent, especially with respect to the
122 taxonomic hierarchical level of study. Hence, characters entailing little or no homoplasy for shallow
123 branches may exhibit greater homoplasy when considered at deeper levels. Goloboff *et al.* (2010)
124 have proposed analytical approaches that take account of this variability in the evolutionary lability
125 of characters along branches.

126 The dominant practice of total evidence analysis for morphology contrasts with the more
127 qualified approach that was sometimes adopted with molecular data (Bull *et al.*, 1993; Felsenstein,
128 1988; Pamilo and Nei, 1988; Maddison, 1997; Nichols, 2001; Degnan and Rosenberg, 2009).

Molecular systematists have debated the relative merits of partitioned versus combined analyses, although the consensus has emerged in favour of the latter (Gatesy *et al.*, 1999; Gatesy and Baker, 2005; Kjer and Honeycutt, 2007; Thompson *et al.*, 2012). Historically, the issue of morphological versus molecular incongruence has been more to the fore (Gatesy and Arctander, 2000), motivated by striking examples of conflict between molecular and morphological cladograms in some groups (Mickevich and Farris, 1981; Bledsoe and Raikow, 1990; Bremer, 1996; Poe, 1996; Baker *et al.,* 1998; Hillis and Wiens, 2000; Wiens and Hollingsworth, 2000; Jenner, 2004; Draper *et al.*, 2007; Springer *et al.*, 2007; Pisani *et al.*, 2007; Near 2003; Mayr, 2011), although see Lee and Camens (2009).

WHY EXAMINE THE CONGRUENCE OF CRANIODENTAL AND POSTCRANIAL PARTITIONS?

Some systematists hold that craniodental and postcranial characters convey signals of differing quality (Ward, 1997; Collard *et al.*, 2001; Naylor and Adams, 2001; Finarelli and Clyde, 2004). However, the evidence for this is piecemeal and largely anecdotal. Many practitioners take a more holistic approach to sampling characters (Sánchez-Villagra and Williams, 1998), sampling densely from as many anatomical regions as possible. However, even where potential characters are reasonably homogenously distributed throughout the body, “certain body regions and organs still hold a considerable mystique for taxonomists as classificatory tools, while others are neglected” (Sokal and Sneath, 1963; page 85). Arratia (2009) noted that actinopterygian systematists focus on cranial characters, despite rich seams of underexploited data within the fin rays and fulcra. Murray and Vickers-Rich (2004) suggested that the crania and mandibles of birds often provide the most informative characters because of their structural complexity. Similarly, Cardini and Elton (2008) demonstrated that characters of the chondrocranium were most informative in studies of *Cercopithecus* monkeys, and suggested that this might apply across primates and perhaps across all mammals. Lastly, Ruta and Bolt (2008) found that characters of the lower jaws of temnospondyls recovered many of the same relationships as those inferred from a more holistic data set.

Studies of extinct organisms *necessarily* focus on those characters capable of fossilisation (typically shells and bones). In vertebrates, disproportionate numbers of characters are often coded from the most recalcitrant skeletal elements, notably teeth in mammals (Billet, 2011; Alejandra Abello, 2013). More generally, there is a tendency for soft part characters to resolve as more derived apomorphies than those characters with a higher preservation potential. Removal of soft characters therefore results in preferential 'stemward slippage' across animal groups (Sansom *et al.,* 2010, 2011; Sansom and Wills, 2013); the phenomenon whereby taxa resolve closer to the root of the tree than they otherwise would.

In this study, we apply a variety of methods to explore differences in the strength and nature of phylogenetic signals in craniodental and postcranial partitions of 85 published vertebrate datasets. We address the following questions: 1. Do levels of homoplasy in craniodental character partitions differ from those in postcranial character partitions (Sanchez-Villagra and Williams, 1998), and are any observed differences more than simply a function of differing numbers of characters within these partitions? We quantify this using conventional indices of homoplasy (Kluge and Farris, 1969; Archie, 1996) modelled with respect to data set parameters. 2. Is there more conflict between craniodental and postcranial characters than we would expect for random partitions, and do craniodental and postcranial characters support significantly different trees as a result? We investigate this using the incongruence length difference test (Mickevich and Farris, 1981; Farris *et al.,* 1995a,b) and a new partition homogeneity test based upon tree distance metrics rather than differences in tree lengths.

Materials and Methods

THE DATA SETS

Phylogenetic datasets published between 2000 and 2013 were sourced from the literature. We restricted our focus to discrete character morphological matrices composed entirely of vertebrate taxa, and analysed using equal weights maximum parsimony. Although morphological data can be

186 analysed with model-based likelihood (Lewis, 2000; Lee and Worthy, 2012) and Bayesian
187 (Nylander *et al*, 2003; Pollitt *et al.*, 2005; Clarke and Middleton, 2008; Tsuj and Mueller, 2009;
188 Bouchard-Cote *et al*, 2012) methods, the considerable majority of published morphological trees
189 are generated utilising maximum parsimony. Matrices were also garnered from Brian O'Meara's
190 *TreeBASE* mirror (O'Meara, 2009), Graeme Lloyd's collection of dinosaur matrices (Lloyd, 2009)
191 and *MorphoBank* (O'Leary and Kaufman, 2011). We excluded matrices with fewer then eight taxa
192 or partitions with fewer than eight parsimony-informative characters (for reasons of statistical
193 power). We interpret craniodental characters here as those pertaining to the skull (cranium plus
194 mandible and dentition).
195 Our resulting sample comprised 85 matrices, spanning all major vertebrate groups. A small
196 number of these datasets contained characters that were not strictly morphological (e.g., character
197 618 relating to habitat choice in the matrix of Spaulding *et al.*, 2009). These were removed prior to
198 any further analysis. We also removed phylogenetically uninformative taxa within partitions using
199 the principles of safe taxonomic reduction (Wilkinson, 1995). We additionally removed taxa with
200 large amounts of missing data that were demonstrated empirically to obfuscate resolution within
201 partitions, usually because they inflated search times and numbers of optimal trees to impractical
202 levels. The number and percentage of craniodental and postcranial characters in each matrix were
203 recorded, as well as the fraction of missing entries within these partitions. Wilcoxon signed-rank
204 tests were used to assess the significance of differences between these medians in different
205 groups.
206 Simply recording percentages of missing cells has limitations, as these can be distributed
207 randomly (usually less problematic analytically) or can be concentrated within particular taxa. Such
208 concentrations are often observed in real data sets, and particularly in matrices including fossil
209 taxa (Cobbett *et al.*, 2007). Our pruning of taxa removed the worst of these effects. Simulations
210 have demonstrated that it is the signal within the characters that *are* coded that is critical in
211 determining the placement of particular taxa (Wiens, 2003ab) rather than numbers of missing cells
212 *per se*.

All phylogenetic analyses were performed using *TNT* (Goloboff *et al.*, 2008b), using equally weighted parsimony. We also reproduced any assumptions regarding character order, polarity and rooting. Empirically, we determined that comprehensive searches involving 200 parsimony ratchet iterations (Nixon, 1999) and 100 drift iterations per replication, with 10 rounds of tree fusion (Goloboff, 1999) were effective at recovering the set of MPTs reported by the original authors in each case. These settings were used throughout.

TESTING WHETHER CRANIODENTAL AND POSTCRANIAL CHARACTERS EXHIBIT DIFFERENT LEVELS OF HOMOPLASY

Consistency and retention indices. – There are two intuitive ways to calculate differences in mean/median consistency indices (ci; Kluge and Farris, 1969) and retention indices (ri; Farris, 1989) for characters in partitions of a dataset (ci and ri in lower case pertain to individual characters). The usual approach is to find the optimal tree or trees for all characters analysed simultaneously (the global MPT(s)) and to take mean values for characters reconstructed on this/these (Sánchez-Villagra and Williams, 1998; Song and Bucheli, 2010). However, there are theoretical partition size effects, even in the absence of differences in the levels of character conflict *within* partitions. All other things being equal, the characters within the larger partition are likely to have higher ci values on average (Fig. 1). The other approach is to report metrics for characters within each partition analysed independently. However, the ensemble CI and ensemble RI (and therefore ci and ri for individual characters) are influenced by data set dimensions (Archie, 1989; Sanderson and Donoghue, 1989; Faith and Cranston, 1991; Klassen *et al.*, 1991): there is a strong, negative correlation between CI and the number of taxa and a weaker, negative relationship between CI and the number of characters (Archie, 1989, 1996; Archie and Felsenstein, 1993).

There are two ways in which differences between indices can be tested. For individual data matrices, Mann-Whitney or t-tests can be applied to character ci and ri values, with the null that these have the same median or mean in the two partitions. For the more general comparison across all 85 matrices simultaneously, Wilcoxon signed ranks or paired t-tests can be used to test

the nulls that (either) the median/mean ci or ri in craniodental and postcranial partitions were similar, or that the median/mean CI and RI indices for the two partitions were similar.

Homoplasy excess ratio (HER). – The homoplasy excess ratio (HER; Archie and Felsenstein, 1993) was proposed as an adjunct to the ensemble consistency index (CI), and argued to be relatively immune to its worst shortcomings. HER is given by (MEANNS – L) / (MEANNS – MINL), where MEANNS is the mean length of the most parsimonious trees resulting from a large sample of matrices (here 999) in which the state assignments within each character have been randomised. L is then the optimal length of the original dataset, and MINL is the minimum possible length of the dataset (number of states minus number of characters). HER was calculated for craniodental and postcranial partitions in isolation, and we then tested for differences in partitions across all 85 datasets using the Wilcoxon signed ranks test.

TESTING INCONGRUENCE BETWEEN CRANIODENTAL AND POSTCRANIAL CHARACTER PARTITIONS

Incongruence length difference (ILD) test. – To assess the significance of congruence between whole character partitions as measured by optimal tree length, the ILD test (Mickevich and Farris, 1981; Farris *et al.,* 1995ab; Barker and Lutzoni, 2002) was applied to the matrices in *TNT* using 999 replicates (Allard *et al.,* 1999a,b). The ILD score is given by $L_{AB} – (L_A + L_B) / L_{AB}$, where L_{AB} is the optimal tree length (in steps) of the simultaneous analysis of both partitions together (the total evidence analysis), and L_A and L_B are the optimal tree lengths for partitions A and B analysed independently. To determine the significance of the observed ILD score, random partitions of the same size (number of characters) as the specified partitions are also generated to yield a distribution of randomized ILD scores. Given the nature of phylogenetic data, the suitability of this test has been questioned on a variety of grounds (Dolphin *et al.* 2000; Hipp *et al.* 2004; Ramirez 2006; Planet 2005, 2006). Despite this, the ILD test remains commonly used to compare the congruence of data partitions. We did not apply the arcsine transformation of Quicke *et al.* (2007) because they justified their correction on the basis of empirical and simulated *molecular* data (morphological data have different statistical properties).

TESTING WHETHER PARTITIONS SUPPORT DIFFERENT TREES

The incongruence relationship difference (IRD) test: a new test of the congruence of relationships. – Much like the ILD test, this is a randomisation-based test. However, partitions are compared via the distances between the optimal trees that result from them, rather than via tree length (ILD) or a matrix-representation of topology (TILD; Wheeler, 1999). There are many possible tree-to-tree distance measures including symmetric difference (RF; Bourque, 1978; Robinson and Foulds, 1981; Pattengale *et al.* 2007) quartets distance (QD; Estabrook *et al.,* 1985), nearest neighbour interchange distance (NNID; Waterman and Smith, 1978), nodal distance (Bluis and Shin, 2003), maximum agreement subtree distance (Goddard *et al.*, 1994; de Vienne *et al.*, 2007), transposition distance (Rossello and Valiente, 2006), subtree prune and regraft distance (SPR; Goloboff, 2008), and path-length difference (PLD; Zaretskii, 1965; Williams and Clifford, 1971). For reasons of familiarity (it is among the most well characterised; e.g. Steel and Penny, 1993) and ease of computation, we chose to use the symmetric difference or Robinson and Foulds distance (RF) (Fig. 2) as our measure of tree-to-tree distance. We note that all other implementations are possible. We illustrate the approach for two examples: firstly the theropod data of Ezcurra and Cuny (2007) (Fig. 3a) and secondly the mammalian data of Beck (2008) (Fig. 3b). The results from each partition are illustrated as 50% majority rule consensus trees for ease of visualization. The left hand tree in both cases is that derived from the analysis of craniodental characters alone, while the right hand tree is inferred from just the postcranial characters. The open circles indicate nodes present in one partition tree that are absent from the other, and the total number of such nodes gives the measure of RF between the trees. This value corresponds to the incongruence relationship difference (IRD_{MR} in the case of the majority rule trees illustrated).

All MPTs from the analysis of each partition were saved and then compared to each other in two different ways. (1) 'Nearest neighbours' (IRD_{NND}) for up to 10,000 trees in each partition: the mean of the minimum distance between each tree in one set, compared with the trees in the other (and *vice versa*) (Cobbett *et al.*, 2007). (2) The distance between the 50% majority-rule consensus trees (from up to 10,000 fundamentals) for each partition (IRD_{MR}) (Figs. 3 and 4). We then

generated random partitions of the original data in the original proportions, and repeated the above exercises in order to yield a distribution of randomized partition tree-to-tree distances. Distances for the original partitions were deemed significantly different from this distribution when they lay in its 5% tail. Our p-values were derived from 999 replicates. All but three (98%) of our 170 partitions yielded less than 10,000 trees; the imposition of a 10,000 bound for the remainder was a necessary limitation to restrict prohibitively long searches in poorly resolved partitions.

Tests not implemented. – The topological incongruence length difference (TILD) test (Wheeler, 1999) is analogous to the ILD test, but is applied to a matrix representation (GIC; Farris, 1973; also known as MRP coding, Baum, 1992; Ragan, 1992) of the branching structure of a consensus of the optimal trees from the data partitions. The test appears to have limited discriminatory power and high type I error rate (Wills *et al.* 2009).

Rodrigo *et al.* (1993) proposed three interrelated tests to investigate differences in relationships directly. The first of these determines whether the symmetrical difference distance (RF: Robinson and Foulds, 1981) between sets of MPTs from independent analyses of the two data set partitions is distinguishable from the distribution of RF distances between a large sample of pairs of random trees. Only weak congruence between partitions is needed to pass this test. The second test of Rodrigo *et al.* (1993) compares the partitions directly, and determines if there is any overlap between the MPTs derived from the two partitions upon bootstrap resampling. This is problematic because the probability of encountering common trees changes with the bootstrap parameters (Lutzoni, 1997), especially the number of replicates (Page, 1996). The third test compares RF distances between partitions with those between trees bootstrapped from within partitions. Although a useful test, it may have limitations, particularly where the partitions of the data set are of very different sizes, and especially where the number of characters in the smaller partition is also small relative to the number of terminals. In such cases, bootstraps of the smaller partition may consistently yield poor resolution and low RF distances between trees *within* this partition (Page, 1996). The IRD_{NND} test proposed above controls this partition size difference.

Results

HOMOPLASY IN CRANIODENTAL AND POSTCRANIAL DATA PARTITIONS

Across our sample of 85 datasets, craniodental partitions had more characters (median = 58) than postcranial partitions (median = 50) (Wilcoxon signed ranks; $V = 2288.5$, $p = 0.044$) (Table 1). The percentage of missing data cells was comparable in craniodental (median = 12.8%) and postcranial (median = 16.9%) partitions, although this difference was significant (paired Wilcoxon; $V = 868.5$, $p = 0.006$). The mean ensemble consistency indices (CI) for craniodental and postcranial characters across all 85 data sets were not significantly different (paired $t = 1.184$, $p = 0.240$), with postcranial partitions ($\bar{X} = 0.564$) having slightly higher values than craniodental partitions ($\bar{X} = 0.550$). Using the mean partition (per character) ci index across all matrices (characters optimised onto the globally optimal tree(s) for the entire matrix) revealed a non-significant difference ($\bar{X} = 0.632$ and 0.627 for craniodental and postcranial partitions respectively; paired $t = 0.450$, $p = 0.654$). Mann-Whitney tests of craniodental and postcranial ci values *within* the 85 data sets yielded 40 significant ($p < 0.05$) results (four or five might be expected). Twenty-one of these 40 had higher means (less homoplasy) for cranial characters, despite their larger partition size. A simple linear model was used to express partition CI in terms of the log of the number of taxa, the log of the number of characters and the log of the percentage of missing data (+1) across all 170 partitions. The term for missing data was not significant, but both the log of the number of taxa ($p < 0.001$) and the log of the number of characters ($p < 0.014$) were highly so (multiple $R^2 = 0.458$, $p = < 0.001$). A subsequent paired t-test of the residual CI values from this model revealed no significant difference ($t = 0.917$, $p = 0.362$) between craniodental and postcranial partitions. We note that other variables have been demonstrated empirically to influence CI (Donoghue and Ree, 2000; Hoyal Cuthill *et al.*, 2010).

Homoplasy excess ratio (HER) values (Archie 1989, 1996; Archie and Felsenstein, 1993) were similar in the craniodental ($\bar{X} = 0.582$) and postcranial skeleton ($\bar{X} = 0.571$) (paired $t = 0.621$, $p = 0.537$). A linear model of HER in terms of the logs of numbers of characters and taxa and the percentage of missing data (+1) revealed no significant independent variables. Finally,

retention indices were significantly higher for postcranial than craniodental partitions when measured across all characters in a partition (RI; paired $t = 2.654$, $p = 0.009$) but not as the average of per character values within a partition (ri; paired $t = 1.538$, $p = 0.128$). Linear modelling of the partition RI in terms of the logs of numbers of characters and taxa and the percentage of missing data (+1) revealed no significant independent variables.

CONGRUENCE BETWEEN CRANIODENTAL AND POSTCRANIAL SIGNALS (ILD TESTS)

When originally described, the ILD test was used with a standard significance level of 5% (0.05). At this level, 31 of our 85 dataset partitions had significant character incongruence (Table 1). Some have advocated more stringent levels (e.g., Cunningham, 1997): with $p < 0.010$, we still rejected the null for 23 of our datasets. We note that the correlation between ILD p-values and the percentage of missing data within a data set was not significant ($\tau = -0.115$, $p = 0.125$), although there was a significant correlation between ILD p-values and the *difference* in the percentage of missing data cells in the two partitions ($\tau = -0.161$, $p = 0.031$). When culling our matrices to those with a difference of just 5% or less in the fraction of missing cells in the two partitions ($n = 40$), we still observed 11 data sets with an ILD significant at $p < 0.050$ (a similar rate of rejection: $G = 2.653$, $p = 0.103$).

Logistic regression was used to model the binary outcome of the ILD test (significant or not at $p < 0.05$) as a function of log of the number of taxa, log of the number of characters, the imbalance in number of characters between partitions (as a percentage of the total number), the percentage of missing entries in the data set, and the log of the imbalance in the percentage of missing entries between partitions. Terms for the number of characters ($p < 0.001$) and the imbalance in character numbers ($p = 0.021$) were retained in the minimum adequate model selected by the progressive deletion of non-significant terms ($p > 0.05$).

THE SIMILARITY OF RELATIONSHIPS IMPLIED BY PARTITIONS

Across all 85 data sets, 27 had significantly incongruent relationships (IRD_{NND}) at $p < 0.05$, of which 14 were also significant at $p < 0.01$ (Table 1). Correlation between p-values and the

difference in the percentage of missing data between partitions was not significant (τ = -0.012, p = 0.876). Moreover, the rate of rejection of the null at $p < 0.05$ was similar for the culled (n = 40) data set and those matrices (n = 45) with more than 5% difference in missing data between partitions (eleven and sixteen significant results respectively) (G = 0.637, p = 0.425). Logistic regression of the binary outcome of the IRD_{NND} test ($p < 0.05$ or otherwise) as above yielded no significant terms in the minimum adequate model.

The correlation between p-values for the nearest neighbour and majority rule variants of the IRD test was highly significant ($p < 0.001$) but not particularly tight (τ = 0.387). The latter offers a very imprecise proxy for the distances measured by nearest neighbours, and we do not advocate its use.

HIGHER TAXONOMIC DIFFERENCES IN CONGRUENCE

Partitioning datasets into six broad and inclusive taxonomic groups (Aves, other Ornithodira, Synapsida, other reptiles, amphibians (including early tetrapods), and fishes) revealed some significant differences (Fig. 5). In particular, synapsids and amphibians were less likely to have congruent partitions according to the ILD test than other groups (G = 11.808, p = 0.038) (Table 2). However, there were significant differences in the log of the number of characters across groups (F = 3.095, p = 0.013), largely accounted for by the contrast between synapsids and fishes (Tukey HSD, p = 0.003). Taxonomic group was not significant as a factor in logistic regression models, and the residuals from models omitting taxonomic group membership retained no differences between groups (Kruskal-Wallis χ-squared = 2.976, p = 0.704). A broadly similar taxonomic pattern was observed for the IRD_{NND} test, although differences in frequencies of null rejection across higher taxa were not significant (G = 4.283, p = 0.510).

Discussion

CRANIODENTAL AND POSTCRANIAL PARTITIONS CONTAIN SIMILAR LEVELS OF HOMOPLASY

417
418 It is well known that the CI is negatively correlated with the number of taxa in a matrix (Archie,
419 1989; Sanderson and Donoghue, 1989; Faith and Cranston, 1991; Klassen *et al.*, 1991), but it also
420 has a weaker, negative relationship with the number of characters (Archie, 1989, 1996; Archie and
421 Felsenstein, 1993). Despite the greater number of craniodental characters than postcranial
422 characters across our sample of data sets, we found no significant difference in CI (paired t =
423 1.184, p = 0.240). Unsurprisingly, when modelling out both data matrix dimensions (numbers of
424 taxa and characters) and the percentage of missing data (+1) across all 170 partitions, residual CI
425 values were even more similar (for residuals: t = 0.917, p = 0.362). Homoplasy excess ratio (HER)
426 values were also similar in the craniodental ($\bar{x}$ = 0.582) and postcranial skeleton ($\bar{x}$ = 0.571)
427 (paired t = 0.621, p = 0.537).
428 We note that the absence of a clear difference between craniodental and postcranial levels
429 of homoplasy does not necessarily imply that *additional* characters of equivalent phylogenetic
430 informativeness can be garnered from the two partitions with comparable ease. One partition may
431 have been exhausted with considerable care, the other not. Our conclusions therefore necessarily
432 relate to the *coded* data. We also note that a high CI within a partition could be the result of a
433 strong phylogenetic signal, or the developmental non-independence of characters. Distinguishing
434 between these causes requires detailed developmental and underpinning genetic knowledge,
435 which are often unavailable.
436
437 **PARTITIONS HAVE INCONGRUENT SIGNALS MORE OFTEN THAN WE EXPECT**
438
439 Our ILD test results demonstrate that our partitions are incongruent about one time in three: 31
440 from 85 datasets. Assuming a significance level (false positive rate) of 5%, we would expect four or
441 five datasets to be significantly incongruent by chance. Significant incongruence is therefore
442 detected across our sample of datasets (binomial test $p < 0.001$, assuming a 5% false positive
443 error rate), although this is partly accounted for by differences in partition parameters. We make no
444 inferences concerning the overall quality of individual data sets on the strength of these results,

445 and note that partitions were imposed by us in each case (rather than reflecting distinctions made
446 by the original authors).
447
448 **CRANIODENTAL AND POSTCRANIAL PARTITIONS OFTEN IMPLY SIGNIFICANTLY**
449 **DIFFERENT RELATIONSHIPS**
450
451 Results from the incongruence relationship difference (IRD_{NND}) tests were broadly similar to
452 those from the ILD test: 32% of data sets yielded significantly different trees from the two partitions.
453 As with the ILD test, we would expect just three or four data sets (5%) to reject the null by chance
454 (a highly significant difference: binomial test $p < 0.001$). Our empirical sample suggests that the
455 IRD_{NND} test using the Robinson Foulds distance is less likely to yield a significant result than the
456 ILD, and is less susceptible to differences in partition size and differences in levels of missing data.
457 The IRD_{NND} therefore has certain advantages over the ILD, and also offers a more intuitive index of
458 congruence (i.e., one based directly on differences in topological branching structure rather than
459 differences in tree length). We note, however, that the IRD_{NND} is insensitive to differences in branch
460 lengths. Many authors have identified difficulties with the incongruence length difference test (ILD)
461 (Dolphin *et al.* 2000; Hipp *et al.* 2004; Ramirez 2006; Planet 2005, 2006) and we do not repeat
462 these here.
463 As with all similar metrics, the IRD is to some extent arbitrary (Wheeler, 1999). Different
464 indices of tree-to-tree distances will yield different distances and p-values, and the IRD can be
465 implemented with many such measures. For example, the Robinson Foulds distance penalises
466 distant and shallow branch transpositions much more strongly than close and shallow
467 transpositions, whereas the maximum agreement subtree distance (Finden and Gordon, 1985), for
468 example, makes only a marginal distinction between these two cases. Similarly, the maximum
469 agreement subtree distance is less sensitive to the depth of the transposition (Cobbett *et al.*,
470 2007). Alternative metrics will have other, and perhaps more desirable, properties (Lin *et al.,* 2011).
471 Our choice of the Robinson Foulds distance here was partly pragmatic, as it is computationally less
472 demanding than most alternatives.
473

WHAT DOES PARTITION INCONGRUENCE IMPLY?

In studying the evolution of form, it is now relatively common to recognize anatomical modules (Mitteroecker and Bookstein, 2007, 2008; Klingenberg, 2008; Cardini and Elton, 2008; Lü *et al.* 2010; Hopkins and Lidgard, 2012; Cardini and Polly, 2013; Goswami *et al.*, 2013, 2015). These are regions of the body (or suites of landmarks) within which morphological changes are strongly correlated through evolutionary time, but between which there is significantly less coordination. Different selective forces may operate on these modules or components of the mosaic (Gould, 1977; Maynard Smith, 1993; Kemp, 2005; Lü *et al.*, 2010), and they may therefore exhibit different evolutionary rates and trends (Mitteroecker and Bookstein, 2007, 2008; Klingenberg, 2008). In the context of phylogenetic characters, differing pressures on modules may favour particular patterns of convergence and homoplasy, and therefore suites of characters that imply different trees (Clarke and Middleton, 2008). The skull of many tetrapod groups has often been regarded as biomechanically and functionally somewhat independent of the rest of the skeleton (Ji *et al.*, 1999; Koski, 2007; Mitteroecker and Bookstein, 2008) hence the difficulty of making many inferences about the one from the other. We note that fishes have the lowest overall incongruence between partitions and the most similar levels of per character consistency index (ci) (Fig. 5). This may result from greater functional and biomechanical integration between the head and trunk in fishes compared with other vertebrate groups (i.e., fishes lack a functional neck) (Klingenberg, 2008; Larouche *et al.*, 2015).

We note that while anatomical modules are usually envisaged as physically proximate suites of landmarks or characters, it is possible for characters to evolve in a coordinated manner across the body as the result of particular selective pressures (Kemp, 2007). For example, adaptations for swimming, digging or flying (Gardiner *et al.*, 2011; Abourachid and Hofling, 2012; Allen *et al.*, 2013) might entail correlated suites of change across the body in a manner that would not be apparent from studying straightforward divisions into body regions (e.g., head, body and limbs). Developmental regulatory processes may also entail counterintuitive suites of coordinated character change across the body (Kharlamova *et al*. 2007; Chase *et al.* 2011), but our objective was not to test for such suites of correlations here.

503 In the most general terms, character selection and coding clearly has an impact upon
504 inferred phylogeny. Most straightforwardly, alternative data sets for identical sets of taxa can yield
505 different trees (Freitas and Brown, 2004; Munoz-Duran, 2011; Penz *et al.* 2013), and this argues
506 strongly for the synthesis of all characters. More generally, systematists rightly exercise their
507 judgement in deciding which aspects of morphology to codify as putative homologies. Wings in
508 birds, bats and pterosaurs are not considered homologous as wings *a priori* (although they are as
509 *limbs*) because the weight of evidence unambiguously rules this out (and in the absence of any
510 formal analysis). In many cases, however, the decision is less straightforward, and the *a priori*
511 omission of characters believed to be analogous or strongly homoplastic may unintentionally
512 overlook useful signal at some level in the tree. Finally, it is often observed empirically that the
513 tree(s) derived from a given matrix can alter markedly with the omission, reweighting or ordering of
514 characters (Wills, 1998), such that even modest perturbations to the data yield large changes to
515 the resulting trees.
516 In morphological phylogenetics, it is usual to combine all available data (Kluge *et al.,* 1989).
517 Our results therefore reinforce the importance of this approach. While the patterns inferred from
518 particular organ systems or suites of characters may be misleading (in the same way and for the
519 same reasons that individual characters may merely introduce homoplasy and noise), combined
520 analysis of all available characters often allows a globally strong phylogenetic signal to emerge
521 from conflicting local homoplasy (Gatesy *et al.* 1999). We demonstrate that in vertebrate studies of
522 this type, an exclusive focus upon characters of either the cranium or postcranium (at the expense
523 of those of the other partition (e.g., Fitzgerald (2010) (craniodental only) and Mayr and Mourer-
524 Chauvire (2004) (postcranial only)) will significantly influence the resultant optimal tree(s) about
525 30% of the time, irrespective of major group. We therefore strongly advocate garnering character
526 data intensively from all anatomical regions whenever possible.
527 When analysing fossils, it is usually impossible to sample across the same suite of
528 characters that would ideally be coded for extant species (Wiens, 2003a,b; Cobbett *et al.,* 2007).
529 For example, in fossil crocodyliforms, the vast majority of characters are coded from the skull (e.g.,
530 O'Connor *et al.,* 2010; Turner and Sertich, 2010; Cau and Fanti, 2011; Hastings *et al.,* 2011;

Puertolas *et al.,* 2011) and it is difficult to be confident that we are not merely inferring a 'craniodental' tree. The only (and indirect) way to test this would be to conduct parallel analyses upon the closest living representatives of the clade. However, the (quite possibly limited) utility of this approach depends upon the phylogenetic proximity of the extant exemplars, the presumed constancy of selective pressures on putative modules through time and across clades (a big assumption: Hunt, 2008; Frazzetta 2012), and the similarity of the available coded data. A related issue in the context of fossil vertebrates is the preferential preservation of hard part characters (bones rather than muscles or other more volatile tissues). An analogous concern, therefore, is whether skeletal and soft-part characters convey a consistent phylogenetic signal (Diogo, 2004). If they do not, then this has implications for the manner in which fossil vertebrates are interpreted and analyzed (Sansom *et al.,* 2010; Sansom and Wills 2013; Pattinson *et al.*, 2014) and is an area particularly needing detailed future work.

Conclusions

1. Systematists typically abstract significantly more characters from the skull than the rest of the skeleton. However, tests for levels of homoplasy in the craniodental and postcranial partitions of our sample of 85 matrices revealed no significant differences, irrespective of how homoplasy was measured. Systematists appear to be coding characters of similar internal consistency from both regions of the body. It is unclear to what extent the bias towards coding craniodental characters reflects a real bias in their distribution, the availability of material, or arises from the preconceptions of systematists.
2. Craniodental and postcranial character partitions exhibited significant incongruence (ILD tests) in 31 of our 85 sample data sets. Likewise, our new IRD_{NND} test found significantly different relationships in 27 from 85 cases.
3. Although vertebrate systematists sometimes code morphological characters preferentially from particular anatomical regions (through necessity or by choice), received wisdom is that a broad and unbiased sampling is usually preferable when possible. Here, we provide

empirical evidence for this view. Our results strongly support dense sampling of characters from across the vertebrate skeleton.

4. The IRD_{NND} test appears to be less sensitive to differences in numbers of characters or amounts of missing data between partitions than the ILD test. We advocate its use as an ancillary test for addressing differences in implied relationships between data partitions.
5. There were significant differences in the distribution of significant partition inhomogeneity across higher taxonomic groups, as measured by the ILD test. Synapsids and amphibians were less likely to have congruent craniodental and postcranial partitions than other groups. However, these differences were no longer significant when the imbalance in partition size was modelled out. Neither were the relationships inferred from craniodental and postcranial characters more likely to conflict significantly in some higher taxa than in others (IRD_{NND} test). Our findings therefore have some generality across vertebrates.

ACKNOWLEDGEMENTS

We would like to thank all authors who kindly supplied us with their datasets (knowingly or otherwise). We are grateful to Marcello Ruta and anonymous reviewers whose comments enabled us to substantially improve this paper. Ward Wheeler and Pablo Goloboff also made invaluable comments on an earlier version of this work as presented at *Hennig XXIX*. Some of our analyses were made possible by the generous free use of a high performance computing cluster; Bioportal, University of Oslo. RCPM gathered data and performed analyses. RS scripted IRD tests within *TNT* and performed analyses. MAW conceived and designed the study, devised/programmed the original tests in *PAUP** and performed analyses. RCPM and MAW summarised results and wrote the paper. We thank Elizabeth Wills for drawing the silhouettes in Figure 3. This work was supported by a University of Bath URS award to RCPM, Leverhulme Trust Grant F/00351/Z and BBSRC grant BB/K015702/1 to MAW, JTF Grant 43915 to Mark Wilkinson and MAW, and NERC fellowship NE/I020253/1 to RSS.

588

589 **LITERATURE CITED**

590

591 Abourachid, A., E. Hofling. 2012. The legs: a key to bird evolutionary success. *J. Ornithol.* **153**: S193-S198.

592 Allain, R., and N. Aquesbi. 2008. Anatomy and phylogenetic relationships of *Tazoudasaurus naimi*
593 (Dinosauria, Sauropoda) from the late Early Jurassic of Morocco. *Geodiversitas* **30**: 345–424.

594 Allain, R., T. Xaisanavong, P. Richir, and B. Khentavong. 2012. The first definitive Asian spinosaurid
595 (Dinosauria: Theropoda) from the Early Cretaceous of Laos. *Naturwissenschaften* **99**: 369-377.

596 Allard, M. W., J. S. Farris, and J. M. Carpenter. 1999a. Congruence among mammalian mitochondrial genes.
597 *Cladistics* **15**: 75–84.

598 Allard, M. W., M. Kearney, K. L. Kivimaki, A. D. Meisner, E. E. Strong. 1999b. The random cladist: A review
599 of the software package *Random Cladistics*. *Cladistics* **15**: 183–189.

600 Allejandra Abello, M. 2013. Analysis of dental homologies and phylogeny of Paucituberculata (Mammalia:
601 Marsupialia). Biol. J. Linn. Soc. 109: 441-465.

602 Allen, V., K.T. Bates, Z.H. Li, J.R. Hutchinson. 2013. Linking the evolution of body shape and locomotor
603 biomechanics in bird-line archosaurs. *Nature* **497**: 104-107.

604 Anderson, J. S., R. R. Reisz, D. Scott, N. B. Frobisch, and S. S. Sumida. 2008. A stem batrachian from the
605 Early Permian of Texas and the origin of frogs and salamanders. *Nature* **453**: 515–518.

606 Andres, B., J. M. Clark, and X. Xing. 2010. A new rhamphorhynchid pterosaur from the Upper Jurassic of
607 Xinjiang China, and the phylogenetic relationships of basal pterosaurs. *J. Vertebr. Palaeontol.* **30**: 163-
608 187.

609 Archie, J. W. 1989. Homoplasy excess ratios: New indices for measuring levels of homoplasy in
610 phylogenetic systematics and a critique of the consistency index. *Syst. Biol.* **38**: 253–269.

611 Archie, J. W. 1996. Measures of homoplasy. Pp. 153-188 *in* M. J. Sanderson, L. Hufford (eds.) *Homoplasy:*
612 *The Recurrence of Similarity in Evolution*. Academic Press, New York.

613 Archie, J. W., and J. Felsenstein. 1993. The number of evolutionary steps on random and minimum length
614 trees for random evolutionary data. *Theor. Popul. Biol.* **43**: 52–79.

615 Arratia, G. 2009. Identifying patterns of diversity of the actinopterygian fulcra. *Acta Zoolog.* **90**: 220–235.

616 Asher, R. J. 2007. A web-database of mammalian morphology and a reanalysis of placental phylogeny. *BMC*
617 *Evol. Biol.* **7**: 108+.

618 Asher, R. J., and M. Hofreiter. 2006. Tenrec phylogeny and the noninvasive extraction of nuclear DNA. *Syst.*
619 *Biol.* **55**: 181–194.

Asher, R. J., and J. Müller. 2012. Molecular tools in palaeobiology: Divergence and mechanisms. Pp. 1-15 *in* R. J. Asher, and J. Müller (eds.) *From Clone to Bone: The Synergy of Morphological and Molecular Tools in Palaeobiology*. Cambridge University Press, Cambridge, UK.

Asher, R.J., J. Meng, J. R. Wible, M. C. McKenna, G. W. Rougier, D. Dashzeveg, M. J. Novacek. 2005. Stem Lagomorpha and the antiquity of Glires. *Science* **307**: 1091–1094.

Baker, R., X. Yu, and R. DeSalle. 1998. Assessing the relative contribution of molecular and morphological characters in simultaneous analysis trees. *Mol. Phylogenet. Evol.* **9**: 427–436.

Barker, F. K., and F. M. Lutzoni. 2002. The utility of the incongruence length difference test. *Syst. Biol.* **51**: 625–37.

Baum, B. R. 1992. Combining trees as a way of combining data sets for phylogenetic inference, and the desirability of combining gene trees. *Taxon* **41**: 3–10.

Beard, K. C., L. Marivaux, Y. Chaimanee, J.-J. Jaeger, B. Marandat, P. Tafforeau, A. N. Soe, S. T. Tun, and A. A. Kyaw. 2009. A new primate from the Eocene Pondaung formation of Myanmar and the monophyly of Burmese amphipithecids. *Proc. Roy. Soc. Lond. B Biol. Sci.* **276**: 3285–3294.

Beck, R. M. D., H. Godthelp, V. Weisbecker, M. Archer, and S. J. Hand. 2008. Australia's oldest marsupial fossils and their biogeographical implications. *PLoS One* **3**: e1858+.

Billet, G. 2011. Phylogeny of the Notoungulata (Mammalia) based on cranial and dental characters. *J. Syst. Palaeontology* **9**: 481-497.

Bledsoe, A. H., and R. J. Raikow. 1990. A quantitative assessment of congruence between molecular and nonmolecular estimates of phylogeny. *J. Mol. Evol.* **30**: 247-259.

Bloch, J. I., M. T. Silcox, D. M. Boyer, and E. J. Sargis. 2007. New paleocene skeletons and the relationship of plesiadapiforms to crown-clade primates. *Proc. Natl. Acad. Sci. Unit. States Am.* **104**: 1159–1164.

Bluis, J., and D. G. Shin. 2003. Nodal distance algorithm: Calculating a phylogenetic tree comparison metric. Pp. 87-94 *in BIBE '03: Proceedings of the 3rd IEEE Symposium on BioInformatics and BioEngineering*. IEEE Computer Society, Washington, DC.

Bouchard-Cote, A., S. Sankararaman, and M. I. Jordan. 2012. Phylogenetic inference via sequential Monte Carlo. *Syst. Biol.* **61**: 579-593.

Bourdon, E., A. Ricqles, and J. Cubo. 2009. A new transantarctic relationship: Morphological evidence for a Rheidae–Dromaiidae–Casuariidae clade (Aves, Palaeognathae, Ratitae). *Zool. J. Linn. Soc.* **156**: 641–662.

650 Bourque, M. 1978. Arbres de steiner et reseaux dont certains sommets sont a localisation variable. Thesis,
651 University of Montreal.
652 Bremer, B. 1996. Combined and separate analyses of morphological and molecular data in the plant family
653 Rubiaceae. *Cladistics* **12**: 21–40.
654 Brochu, C. A., J. Njau, R. J. Blumenschine, and L. D. Densmore. 2010. A new horned crocodile from the
655 Plio-Pleistocene hominid sites at Olduvai Gorge, Tanzania. *PLoS One* **5**: e9333.
656 Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and
657 combining data in phylogenetic analysis. *Syst. Biol.* **42**: 384–397.
658 Burns, M. E., P. J. Currie, R. L. Sissons, and V. M. Arbour. 2011. Juvenile specimens of *Pinacosaurus*
659 *grangeri* Gilmore, 1933 (Ornithischia: Ankylosauria) from the Late Cretaceous of China, with
660 comments on the specific taxonomy of Pinacosaurus. *Cretaceous Research* **32**: 174-186.
661 Butler, R.J., P. Upchurch, and D. B. Norman. 2008. The phylogeny of the Ornithischian dinosaurs. *J. Syst.*
662 *Palaeontology* **6**: 1–40.
663 Cardini, A., and S. Elton. 2008. Does the skull carry a phylogenetic signal? Evolution and modularity in the
664 guenons. *Biol. J. Linn. Soc.* **93**: 813–834.
665 Cardini, A., and P.D. Polly. 2013. Larger mammals have longer faces because of size-related constraints on
666 skull form. *Nature Comms.* **4**: 2458.
667 Carrano, M.T., and S. D. Sampson. 2008. The phylogeny of Ceratosauria (Dinosauria: Theropoda). *J. Syst.*
668 *Palaeontology* **6**: 183–236.
669 Cau, A., and F. Fanti. 2011. The oldest known metriorhynchid crocodylian from the middle Jurassic of north-
670 eastern Italy: *Neptunidraco ammoniticus* gen. et sp. nov. *Gondwana Research* **19**: 550-565.
671 Chase, K., D.R. Carrier, F.R. Adler, T. Jarvik, E.A. Ostrander, T.D. Lorentzen, and K.G. Lark. 2002. Genetic
672 basis for systems of skeletal quantitative traits: Principal component analysis of the canid skeleton.
673 *Proc. Natl. Acad. Sci. Unit. States. Am.* **99**: 9930-9935.
674 Cheng, L., X.-H. Chen, X.-W. Zeng, and Y.- J. Cai. 2012. A new eosauropterygian (Diapsida: Sauropterygia)
675 from the Middle Triassic of Luoping, Yunnan Province. *J. Earth Sci.* **23**: 33-40.
676 Choo, B. 2011. Revision of the actinopterygian genus *Mimipiscis* (= *Mimia*) from the Upper Devonian Gogo
677 Formation of Western Australia and the interrelationships of the early Actinopterygii. *Trans. Earth. Sci.*
678 **102**: 77-104.

679 Clarke, J. A., and K. M. Middleton. 2008. Mosaicism, modules, and the evolution of birds: Results from a
680 Bayesian approach to the study of morphological evolution using discrete character data. *Syst. Biol.*
681 **57**: 185-201.
682 Cobbett A., M. Wilkinson, and M. A. Wills. 2007. Fossils impact as hard as living taxa in parsimony analyses
683 of morphology. *Syst. Biol.* **56**: 753-766.
684 Collard, M., S. Gibbs, and B. Wood. 2001. Phylogenetic utility of higher primate postcranial morphology. *Am.*
685 *J. Phys. Anthropol. Suppl.* **32**: 52.
686 Cunningham, C. W. 1997. Can three incongruence tests predict when data should be combined? *Mol. Biol.*
687 *Evol.* **14**: 733–740.
688 de Pietri, V.L., L. Costeur, L. Guntert, and G. Mayr. 2011. A revision of the Lari (Aves, Charadriiformes) from
689 the early Miocene of Saint-Gerand-le-Puy (Allier, France). *J. Vertebr. Paleontol.* **31**: 812-828.
690 de Vienne, D. M., T. Giraud, O. C. Martin. 2007. A congruence index for testing topological similarity
691 between trees. *Bioinformatics (Oxford, England)* **23**: 3119–3124.
692 Degnan, J. H., and N. A. Rosenberg. 2009. Gene tree discordance, phylogenetic inference and the
693 multispecies coalescent. *Trends Ecol. Evol*. **24**: 332–340.
694 Diogo, R. 2004. *Morphological Evolution, Adaptations, Homoplasies, Constraints, and Evolutionary Trends:*
695 *Catfishes as a Case Study on General Phylogeny and Macroevolution*. Science Publishers, Enfield.
696 Dolphin, K., R. Belshaw, C. D. L. Orme, and D. L. J. Quicke. 2000. Noise and incongruence: Interpreting
697 results of the incongruence length difference test. *Mol. Phylogenet. Evol.* **17**: 401–406.
698 Donoghue, M. J., and R. H. Ree. 2000. Homoplasy and developmental constraint: A model and an example
699 from plants. *Am. Zool.* **40**: 759-769.
700 Draper, I., L. Hedenas, and G. L. Grimm. 2007. Molecular and morphological incongruence in European
701 species of *Isothecium* (Bryophyta). *Mol. Phylogenet. Evol.* **42**: 700–716.
702 Dutel, H., J. G. Maisey, D. R. Schwimmer, P. Janvier, M. Herbin, and G. Clement. 2012. The giant
703 Cretaceous coelacanth (Actinistia, Sarcopterygii) *Megalocoelacanthus dobiei* Schwimmer, Stewart
704 and Williams, 1994, and its bearing on Latimerioidei interrelationships. *PLoS One* **7**: DOI:
705 10.1371/journal.pone.0049911
706 Estabrook, G.F., F. R. McMorris, and C. A. Meacham. 1985. Comparison of undirected trees based on
707 subtrees of four evolutionary units. *Syst. Zool.* **34**: 193–200.

Ezcurra, M. D., and G. Cuny. 2007. The coelophysoid *Lophostropheus airelensis*, gen. nov.: A review of the systematics of "*Liliensternus*" airelensis from the Triassic-Jurassic outcrops of Normandy (France). *J. Vertebr. Paleontol.* **27**: 73–86.

Faith, D. P., and P. S. Cranston. 1991. Could a cladogram this short have arisen by chance alone?: On permutation tests for cladistic structure. *Cladistics* **7**: 1–28.

Farke, A. A., M. J. Ryan, P. M. Barrett, D. H. Tanke, D. R. Braman, M. A. Loewen, and M. R. Graham. 2011. A new centrosaurine from the Late Cretaceous of Alberta, Canada, and the evolution of parietal ornamentation in horned dinosaurs. *Acta Palaeontologica Polonica* **56**: 691-702.

Farris, J. S. 1973. On comparing the shapes of taxonomic trees. *Syst. Zool.* **22**: 50–54.

Farris, J. S. 1989. The retention index and homoplasy excess. *Syst. Biol.* **38**: 406–407.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995a. Constructing a significance test for incongruence. *Syst. Biol.* **44**: 570–572.

Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1995b. Testing significance of incongruence. *Cladistics* **10**: 315–319.

Felsenstein, J. 1988. Phylogenies from molecular sequences: Inference and reliability. *Annu. Rev. Genet.* **22**: 521–565.

Finarelli, J. A., and W. C. Clyde. 2004. Reassessing hominoid phylogeny: Evaluating congruence in the morphological and temporal data. *Paleobiology* **30**: 614-651.

Finden, C.R. and A.D. Gordon. 1985. Obtaining common pruned trees. *J. Classif.* **2**: 255-276.

Fitzgerald, E. M. G. 2010. The morphology and systematics of *Mammalodon colliveri* (Cetacea: Mysticeti), a toothed mysticete from the Oligocene of Australia. *Zool. J. Linn. Soc.* **158**: 367-476.

Frazzetta, T. H. 2012. Flatfishes, turtles, and bolyerine snakes: Evolution by small steps or large, or both? *Evol. Biol.* **39**: 30-60.

Freitas, A. V. L., and K. S. Brown. 2004. Phylogeny of the Nymphalidae (Lepidoptera). *Syst. Biol.* **53**: 363-383.

Friedman, M. 2007. *Styloichthys* as the oldest coelacanth: Implications for early osteichthyan interrelationships. *J. Syst. Palaeontology* **5**: 289–343.

Friedman, M. 2008. The evolutionary origin of flatfish asymmetry. *Nature* **454**: 209–212.

Frobisch, J. 2007. The cranial anatomy of *Kombuisia frerensis* Hotton (Synapsida, Dicynodontia) and a new phylogeny of anomodont therapsids. *Zool. J. Linn. Soc.* **150**: 117–144.

738 Gaffney, E. S., G. E. Hooks, and V.P. Schneider. 2009. New material of North American side-necked turtles
739 (Pleurodira: Bothremydidae). *American Museum Novitates* **3655**: 1–26.
740 Gardiner, J.D., J.R. Codd, R.L. Nudds. 2011. An association between ear and tail morphologies of bats and
741 their foraging style. *Can. J. Zool.* **89**: 90-99.
742 Gates, T.A., and S. D. Sampson. 2007. A new species of *Gryposaurus* (Dinosauria: Hadrosauridae) from the
743 late Campanian Kaiparowits formation, Southern Utah, USA. *Zool. J. Linn. Soc.* **151**: 351–376.
744 Gatesy, J., and P. Arctander. 2000. Hidden morphological support for the phylogenetic placement of
745 *Pseudoryx nghetinhensis* with bovine bovids: A combined analysis of gross anatomical evidence and
746 DNA sequences from five genes. *Syst. Biol.* **49**: 515-538.
747 Gatesy, J., and R.H. Baker. 2005. Hidden likelihood support in genomic data: Can forty-five wrongs make a
748 right? *Syst. Biol.* **54**: 483-492.
749 Gatesy, J., P. O'Grady, and R. H. Baker. 1999. Corroboration among data sets in simultaneous analysis:
750 Hidden support for phylogenetic relationships among higher level artiodactyl taxa. *Cladistics* **15**: 271–
751 313.
752 Gaubert, P., W. C. Wozencraft, P. Cordeiro-Estrela, and G. Veron. 2005. Mosaics of convergences and
753 noise in morphological phylogenies: What's in a viverrid-like carnivoran? *Syst. Biol.* **54**: 865–894.
754 Gaudin, T., R. Emry, and J. Wible. 2009. The phylogeny of living and extinct pangolins (Mammalia,
755 Pholidota) and associated taxa: A morphology based analysis. *J. Mamm. Evol.* **16**: 235–305.
756 Goddard, W., E. Kubicka, G. Kubicki, and F. R. McMorris. 1994. The agreement metric for labelled binary
757 trees. *Math. Biosci.* **123**: 215–226.
758 Godefroit, P., H. Shulin, Y. Tingxiang, and P. Lauters. 2008. New hadrosaurid dinosaurs from the uppermost
759 Cretaceous of Northeastern China. *Acta Palaeontologica Polonica* **53**: 47–74.
760 Goloboff, P. A. 1999. Analyzing large data sets in reasonable times: Solutions for composite optima.
761 *Cladistics* **15**: 415-428.
762 Goloboff, P. A. 2010. On weighting characters differently in different parts of the cladogram. *Cladistics* **26**:
763 212.
764 Goloboff, P. A. 2008. Calculating SPR distances between trees. *Cladistics* **24**: 591–597.
765 Goloboff, P. A. 2014. Extended implied weighting. *Cladistics* **30**: 260–272.
766 Goloboff, P. A., J. M. Carpenter, J. S. Arias, and D. R. M. Esquivel. 2008a. Weighting against homoplasy
767 improves phylogenetic analysis of morphological data sets. *Cladistics* **24**: 758-773.

768 Goloboff, P.A., J. S. Farris, and K. C. Nixon. 2008b. TNT, a free program for phylogenetic analysis. *Cladistics*
769 **24**: 774-786.
770 Goswami, A., N. Milne, and S. Wroe. 2011. Biting through constraints: cranial morphology, disparity and
771 convergence across living and fossil carnivorous mammals. *Proc. Roy. Soc. Lond. B Biol. Sci.* **278**:
772 1831-1839.
773 Goswami, A., W.J. Binder, J. Meachen, and F.R. O'Keefe. 2015. The fossil record of phenotypic integration
774 and modularity: A deep-time perspective on developmental and evolutionary dynamics. *Proc. Natl.*
775 *Acad. Sci. Unit. States. Am.* **112**: 4891-4896.
776 Gould, G. C. 2001. The phylogenetic resolving power of discrete dental morphology among extant
777 hedgehogs and the implications for their fossil record. *American Museum Novitates* **3340**: 1-52.
778 Gould, S. J. 1977. *Ontogeny and Phylogeny*. Belknap Press of Harvard University Press, Harvard.
779 Hamrick, M. W. 2012. The developmental origins of mosaic evolution in the primate limb skeleton. *Evol. Biol.*
780 **39**: 447-455.
781 Hastings, A. K., J. I. Bloch, and C. A. Jaramillo. 2011. A new longirostrine dyrosaurid (Crocodylomorpha,
782 Mesoeucrocodylia) from the Paleocene of north-eastern Colombia: Biogeographic and behavioural
783 implications for New-World Dyrosauridae. *Palaeontology* **54**: 1095-1116.
784 Hill, R. V. 2005. Integration of morphological data sets for phylogenetic analysis of Amniota: The importance
785 of integumentary characters and increased taxonomic sampling. *Syst. Biol*. **54**: 530–547.
786 Hillis, D. M., and J. J. Wiens. 2000. Molecules versus morphology in systematics: Conflicts, artifacts, and
787 misconceptions. Pp. 1-19 in J. J. Wiens (ed.) *Phylogenetic Analysis of Morphological Data*.
788 Smithsonian Institution Press, Washington DC.
789 Hilton, E. J., and P. L. Forey. 2009. Redescription of ?*Chondrosteus acipenseroides* Egerton, 1858
790 (Acipenseriformes, ?Chondrosteidae) from the lower Lias of Lyme Regis (Dorset, England), with
791 comments on the early evolution of sturgeons and paddlefishes. *J. Syst. Palaeontology* **7**: 427–453.
792 Hipp, A. L., J. C. Hall, and K. J. Sytsma, 2004. Congruence versus phylogenetic accuracy: Revisiting the
793 incongruence length difference test. *Syst. Biol.* **53**: 81–89.
794 Holland, T., and J. A. Long. 2009. On the phylogenetic position of *Gogonasus andrewsae*, within the
795 Tetrapodomorpha. *Acta Zoolog.* **90**: 285-296.
796 Hopkins, M. J, and S. Lidgard. 2012. Evolutionary mode routinely varies among morphological traits within
797 fossil species lineages. *Proc. Natl. Acad. Sci. Unit. States. Am.* **50**: 20520-20525.

798 Hospitaleche, C. A., C. Tambussi, M. Donato, and M. Cozzuol. 2007. A new Miocene penguin from
799 Patagonia and its phylogenetic relationships. *Acta Palaeontologica Polonica* **52**: 299–314.
800 Hoyal-Cuthill, J.F., S. J. Braddy, and P. C. J. Donoghue. 2010. A formula for maximum possible steps in
801 multistate characters: Isolating matrix parameter effects on measures of evolutionary convergence.
802 *Cladistics* **26**: 98-102.
803 Hunt, G. 2008. Gradual or pulsed evolution: When should punctuational explanations be preferred?
804 *Paleobiology* **34**: 360-377.
805 Hurley, I. A., R. L. Mueller, K. A. Dunn, E. J. Schmidt, M. Friedman, R. K. Ho, V. E. Prince, Z. Yang, M. G.
806 Thomas, and M. I. Coates. 2007. A new time-scale for ray-finned fish evolution. *Proc. Roy. Soc. Lond.*
807 *B Biol. Sci.* **274**: 489–498.
808 Huson, D. H., and C Scornavacca. 2012. Dendroscope 3: An interactive tool for rooted phylogenetic trees
809 and networks. *Syst. Biol.* **61**: 1061-1067.
810 Imamura, H., S. M. Shirai, and M. Yabe. 2005. Phylogenetic position of the family Trichodontidae (Teleostei:
811 Perciformes), with a revised classification of the perciform suborder Cottoidei. *Ichthyological Research*
812 **52**: 264–274.
813 Jenner, R. A. 2004. When molecules and morphology clash: Reconciling conflicting phylogenies of the
814 Metazoa by considering secondary character loss. *Evol. Dev.* **6**: 372–378.
815 Ji, Q., Z.-X. Luo, and S.-A. Ji. 1999. A Chinese triconodon mammal and mosaic evolution of the mammalian
816 skeleton. *Nature* **398**: 326-330.
817 Ji, C., D.-Y. Jiang, R. Motani, W.-C. Hao, Z.-Y. Sun, and T. Cai. 2013. A new juvenile specimen of
818 *Guanlingsaurus* (Ichthyosauria, Shastasauridae) from the Upper Triassic of southwestern China. *J.*
819 *Vertebr. Paleontol.* **33**: 340-348.
820 Kangas, A. T., A. R. Evans, I. Thesleff, and J. Jernvall. 2004. Nonindependence of mammalian dental
821 characters. *Nature* **432**, 211-214.
822 Kemp, T. S. 2005. *The Origin and Evolution of Mammals*. Oxford University Press, Oxford.
823 Kemp, T. S. 2007. The concept of correlated progression as the basis of a model for the evolutionary origin
824 of major new taxa. *Proc. Roy. Soc. Lond. B Biol. Sci.* **274**: 1667-1673.
825 Kharlamova, A.V., L.N. Trut, D.R. Carrier, K. Chase and K.G. Lark. 2007. Genetic regulation of canine
826 skeletal traits: trade-offs between the hind limbs and forelimbs in the fox and dog. *Integr. Comp. Biol.*
827 **47**: 373–381.

828 Kitching, I., P. Forey, C. Humphries, and D. Williams. 1998. *Cladistics: Theory and Practice of Parsimony*
829 *Analysis (Systematics Association Special Volume)*. Oxford University Press, Oxford.
830 Kivell, T. L., A. P. Barros, and J. B. Smaers. 2013. Different evolutionary pathways underlie the morphology
831 of wrist bones in hominoids. *BMC Evol. Biol.* **13**: 1-12.
832 Kjaer. K. M., and R. L. Honeycutt. 2007. Site specific rates of mitochondrial genomes and the phylogeny of
833 Eutheria. *BMC Evol. Biol.* **7**: 8.
834 Klassen, G. J., R. D. Mooi, and A. Locke. 1991. Consistency indices and random data. *Syst. Zool.* **40**: 446-
835 457.
836 Klingenberg, C. P. 2008. Morphological integration and developmental modularity. *Annu. Rev. Ecol. Evol.*
837 *Systemat.* **39**: 115-132.
838 Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates*
839 (Boidae, Serpentes). *Syst. Zool.* **38**: 7–25.
840 Kluge, A. G., and J. S. Farris. 1969. Quantitative phyletics and the evolution of anurans. *Syst. Zool.* **18**: 1–
841 32.
842 Koski, K. 2007. The mandibular complex. *Eur. J. Orthod.* **29**: i118-i123.
843 Larouche, O., R. Cloutier, M.L Zelditch. 2015. Head, body and fins: patterns of morphological integration and
844 modularity in fishes. *Evol. Biol.* **42**: 296-311.
845 Lee, M. S. Y., and A. B. Camens. 2009. Strong morphological support for the molecular evolutionary tree of
846 placental mammals. *J. Evol. Biol.* **22**: 2243-2257.
847 Lee, M. S. Y., and J. D. Scanlon. 2002. Snake phylogeny based on osteology, soft anatomy and ecology.
848 *Biol. Rev.* **77**: 333–401.
849 Lee, M. S. Y., and T. H. Worthy. 2012. Likelihood reinstates *Archaeopteryx* as a primitive bird. *Biol. Lett.* **8**:
850 299-303.
851 Leighton, L. R., and C. G. Maples. 2002. Evaluating internal versus external characters: Phylogenetic
852 analyses of the Echinoconchidae, Buxtoniinae, and Juresaniinae (Phylum Brachiopoda). *J. Paleontol.*
853 **76**: 659–671.
854 Le Quesne, W. J. 1969. A method of selection of characters in numerical taxonomy. *Syst. Biol.* **18**: 201–205.
855 Lewis, P. 2000. A likelihood approach to estimating phylogeny from discrete morphological character data.
856 *Syst. Biol.* **50**: 913-925.
857 Li, P.-P., K.-Q. Gao, L.-H. Hou, and X. Xu. 2007. A gliding lizard from the early Cretaceous of China. *Proc.*
858 *Natl. Acad. Sci. Unit. States Am.* **104**: 5507–5509.

859 Lin, Y., V. Rajan, and B. Moret. 2011. A metric for phylogenetic trees based on matching bioinformatics
860 research and applications. *Lect. Notes Comput. Sci.* **6674**: 197-208
861 Lister, A. M., C. J. Edwards, D. A. W. Nock, M. Bunce, I. A. van Pijlen, D. G. Bradley, M. G. Thomas, and I.
862 Barnes. 2005. The phylogenetic position of the 'giant deer' *Megaloceros giganteus*. *Nature* **438**: 850–
863 853.
864 Lloyd, G. T. 2009. Grame T. Lloyd Matrices. http://www.graemetlloyd.com/matr.html.
865 Lopez-Arbarello, A. 2011. Phylogenetic interrelationships of ginglymodian fishes (Actinopterygii:
866 Neopterygii). *PLoS One* **7**: e39370. doi:10.1371/journal.pone.0039370.
867 Lü, J.-C., D. M. Unwin, X. Jin, Y. Liu, and Q. Ji. 2010. Evidence for modular evolution in a long-tailed
868 pterosaur with a pterodactyloid skull. *Proc. Roy. Soc. Lond. B Biol. Sci.* **277**: 383–389.
869 Lü, J.-C., P. J. Currie, L. Xu, X.-L. Zhang, H.-Y. Pu, and S.-H. Jia. 2013. Chicken-sized oviraptorid dinosaurs
870 from central China and their ontogenetic implications. *Naturwissenschaften* **100**: 165-175.
871 Lyson, T. R., and W. G. Joyce. 2009. A new species of *Palatobaena* (Testudines: Baenidae) and a maximum
872 parsimony and Bayesian phylogenetic analysis of Baenidae. *J. Paleontol.* **83**, 457–470.
873 Maddison, W. P. 1997. Gene trees in species trees. *Syst. Biol.* **46**: 523–536.
874 Manegold, A., and T. Toepfer. 2013. The systematic position of *Hemicircus* and the stepwise evolution of
875 adaptations for drilling, tapping and climbing up in true woodpeckers (Picinae, Picidae). *J. Zool.*
876 *Systemat. Evol. Res.* **51**: 72-82.
877 Martinelli, A. G., and G. W. Rougier. 2007. On *Chaliminia musteloides* (Eucynodontia: Tritheledontidae) from
878 the Late Triassic of Argentina, and a phylogeny of Ictidosauria. *J. Vertebr. Paleontol.* **27**: 442–460.
879 Martinez, R. N., and O. A. Alcober. 2009. A basal sauropodomorph (Dinosauria: Saurischia) from the
880 Ischigualasto formation (Triassic, Carnian) and the early evolution of Sauropodomorpha. *PLoS One* **4**:
881 e4397+.
882 Matsumoto, R., S. Suzuki, K. Tsogtbaatar, and S. Evans. 2009. New material of the enigmatic reptile
883 *Khurendukhosaurus* (Diapsida: Choristodera) from Mongolia. *Naturwissenschaften* **96**: 233–242.
884 Maurício, G. N., J. I. Areta, M. R. Bornschein, and R. E. Reis. 2012. Morphology-based phylogenetic
885 analysis and classification of the family Rhinocryptidae (Aves: Passeriformes). *Zool. J. Linn. Soc.* **166**:
886 377-432.
887 Maynard Smith, J. 1993. *The Theory of Evolution*. Cambridge University Press, Cambridge.
888 Mayr, G. 2009. Phylogenetic relationships of the paraphyletic caprimulgiform birds (nightjars and allies). *J.*
889 *Zool. Syst. Evol. Res.* **48**: 126–137.

Mayr, G., R. S. Rana, K. D. Rose, A. Sahni, K. Kumar, L. Singh, and T. Smith. 2010. Quercypsitta-like birds from the early Eocene of India (Aves, ?Psittaciformes). *J. Vertebr. Paleontol.* **30**: 467-478.

Mayr, G. 2011. The phylogeny of charadriiform birds (shorebirds and allies) – reassessing the conflict between morphology and molecules. *Zool. J. Linn. Soc.* **161**: 916–934.

Mayr, G., and C. Mourer-Chauviré. 2004. Unusual tarsometatarsus of a mousebird from the Paleogene of France and the relationships of *Selmes* Peters, 1999. *J. Vertebr. Paleontol.* **24**: 366-372.

Mickevich, M. F. 1978. Taxonomic congruence. *Syst. Biol.* **27**: 143–158.

Mickevich, M. F., and J. S. Farris. 1981. The implications of congruence in *Menidia*. *Syst. Zool.* **30**: 351-370.

Mitteroecker, P., and F. Bookstein. 2007. The conceptual and statistical relationship between modularity and morphological integration. *Syst. Biol.* **56**: 818-836.

Mitteroecker, P., and F. Bookstein. 2008. The evolutionary role of modularity and integration in the hominoid cranium. *Evolution* **62**: 943-958.

Muller, J., and R. R. Reisz. 2006. The phylogeny of early eureptiles: Comparing parsimony and Bayesian approaches in the investigation of a basal fossil clade. *Syst. Biol.* **55**: 503–511.

Munoz-Duran, J. 2011. Data set incongruence, misleading characters, and insights from the fossil record: The canid phylogeny. *Caldasia* **33**: 637-658.

Murray, P.F., and P. Vickers-Rich. 2004. *Magnificent Mihirungs: The Colossal Flightless Birds of the Australian Dreamtime*. Indiana University Press, Bloomington.

Naylor, G. J. P., and D. C. Adams. 2001. Are the fossil data really at odds with the molecular data? Morphological evidence for Cetartiodactyla phylogeny reexamined. *Syst. Biol.* **50**: 444-453.

Nesbitt, S. J., N. D. Smith, R. B. Irmis, A. H. Turner, A. Downs, and M. A. Norell. 2009. A complete skeleton of a Late Triassic saurischian and the early evolution of dinosaurs. *Science* **326**: 1530-1533.

Nesbitt, S. J. 2011. The early evolution of archosaurs: Relationships and the origin of major clades. *Bull. Am. Mus. Nat. Hist.* **352**: 1-292.

Nesbitt, S. J., D. T. Ksepka, and J. A. Clarke . 2011. Podargiform affinities of the enigmatic *Fluvioviridavis platyrhamphus* and the early diversification of strisores ("Caprimulgiformes" + Apodiformes). *PLoS One* **6**: e26350.

Nichols, R. 2001. Gene trees and species trees are not the same. *Trends Ecol. Evol.* **16**: 358–364.

Nixon, K. C. 1999. The parsimony ratchet, a new method for rapid parsimony analysis. *Cladistics* **15**: 407-414.

Nylander, J. A. A, F. Ronquist, J. P. Huelsenbeck, and J. L. Nieves-Aldrey. 2003. Bayesian phylogenetic analysis of combined data. *Syst. Biol.* **53**: 47-67.

O'Connor, J.K., X. Wang, L.M. Chiappe, C. Gao, Q. Meng. 2009. Phylogenetic support for a specialized clade of Cretaceous Enantiornithine birds with information from a new species. *J. Vert. Paleontol*. **29**: 188–204.

O'Connor, P. M., J. J. W. Sertich, N. J. Stevens, E. M. Roberts, M. D. Gottfried, T. L. Hieronymus, Z. A. Jinnah, R. Ridgely, S. E. Ngasala, and J. Temba. 2010. The evolution of mammal-like crocodyliforms in the Cretaceous period of Gondwana. *Nature* **466**: 748-751.

O'Leary, M.A., J. Gatesy, and M. J. Novacek. 2003. Are the dental data really at odds with the molecular data? Morphological evidence for whale phylogeny (Re)Reexamined. *Syst. Biol.* **52**: 853–864.

O'Leary, M. A., and S. Kaufman. 2011. MorphoBank: Phylophenomics in the "cloud". *Cladistics* **27**: 529-537.

O'Meara, B. 2009. Brian O'Meara Lab. http://www.brianomeara.info/treebase.

Osi, A., and L. Makadi. 2009. New remains of *Hungarosaurus tormai* (Ankylosauria, Dinosauria) from the Upper Cretaceous of Hungary: Skeletal reconstruction and body mass estimation. *Paleontologische Zeitschrift* **83**: 227–245.

Page, R. D. 1996. On consensus, confidence, and 'total evidence'. *Cladistics* **12**: 83–92.

Palci, A., M. W. Caldwell, and A. M. Albino. 2013. Emended diagnosis and phylogenetic relationships of the Upper Cretaceous fossil snake *Najash rionegrina* Apesteguía and Zaher, 2006. *Journal of Vertebrate Paleontology* **33**: 131-140.

Pamilo, P., M. Nei. 1988. Relationships between gene trees and species trees. *Mol. Biol. Evol.* **5**: 568–583.

Parenti, L. R. 2008. A phylogenetic analysis and taxonomic revision of ricefishes, *Oryzias* and relatives (Beloniformes, Adrianichthyidae). *Zool. J. Linn. Soc.* **154**: 494–610.

Parrilla-Bel, J., M. T. Young, M. Moreno-Azanza, and J. I. Canudo. 2013. The first metriorhynchid crocodylomorph from the Middle Jurassic of Spain, with implications for evolution of the subclade Rhacheosaurini. *PLoS One* **8**: e54275+

Pattengale, N. D., E. J. Gottlieb, and B. M. E. Moret. 2007. Efficiently computing the Robinson-Foulds metric. *J. Comput. Biol.* **14**: 724–735.

Pattinson, D. J., R. S. Thompson, A. Piotrowski, and R. J. Asher. 2014. Phylogeny, palaeontology, and primates: Do incomplete fossils bias the tree of life? *Syst. Biol.* **64**: 169-86.

949 Penz, C. M., A. V. L. Freitas, L. A. Kaminski, M. M. Casagrande, and P. J. Devries. 2013. Adult and early-
950 stage characters of Brassolini contain conflicting phylogenetic signal (Lepidoptera, Nymphalidae).
951 *Syst. Entomol.* **38**: 316-333.
952 Phillips, M. J., T. H. Bennett, and M. S. Y. Lee. 2009. Molecules, morphology, and ecology indicate a recent,
953 amphibious ancestry for echidnas. *Proc. Natl. Acad. Sci. Unit. States Am.* **106**: 17089–17094.
954 Pine, R. H., R. M. Timm, and M. Weksler. 2012. A newly recognized clade of trans-Andean oryzomyini
955 (rodentia: Cricetidae), with description of a new genus. *J. Mammal.* **93**: 851-870.
956 Pisani, D., M. J. Benton, and M. Wilkinson. 2007. Congruence of morphological and molecular phylogenies.
957 *Acta Biotheor.* **55**: 269-281.
958 Planet, P. 2006. Tree disagreement: Measuring and testing incongruence in phylogenies. *J. Biomed.*
959 *Informat.* **39**: 86–102.
960 Planet, P. J., and I. N. Sarkarm. 2005. mILD: A tool for constructing and analyzing matrices of pairwise
961 phylogenetic character incongruence tests. *Bioinformatics* **21**: 4423–4424.
962 Poe, S. 1996. Data set incongruence and the phylogeny of crocodilians. *Syst. Biol.* **45**: 393–414.
963 Pollitt, J. R., R. A. Fortey and M. A. Wills. 2015. Systematics of the trilobite families Lichidae Hawle & Corda,
964 1847 and Lichakephalidae Tripp, 1957: The application of Bayesian inference to morphological data.
965 *J. Syst. Palaeontology* **3**: 225-241.
966 Poyato-Ariza, F. J. 2003. Dental characters and phylogeny of pycnodontiform fishes. *J. Vertebr. Paleontol.*
967 **23**: 937-940.
968 Puértolas, E., J. I. Canudo, and P. Cruzado-Caballero. 2011. A new crocodylian from the late Maastrichtian
969 of Spain: Implications for the initial radiation of crocodyloids. *PLoS One* **6**: e20011+.
970 Pujos, F., G. de Iuliis, C. Argot, and L. Werdelin. 2007. A peculiar climbing Megalonychidae from the
971 Pleistocene of Peru and its implication for sloth history. *Zool. J. Linn. Soc.* **149**: 179–235.
972 Quicke, D. L., O. R. Jones, and D. R. Epstein. 2007. Correcting the problem of false incongruence due to
973 noise imbalance in the incongruence length difference (ILD) test. *Syst. Biol.* **56**: 496–503.
974 Rae, T. C. 1999. Mosaic evolution in the origin of the Hominoidea. *Folia Primatol.* **70**: 125-135.
975 Ragan, M. 1992. Matrix representation in reconstructing phylogenetic relationships among the eukaryotes.
976 *Biosystems* **28**: 47–55.
977 Ramirez, M. J. 2006. Further problems with the incongruence length difference test: “Hypercongruence”
978 effect and multiple comparisons. *Cladistics* **22**: 289–295.
979 Robinson, D. R., and L. R. Foulds. 1981. Comparison of phylogenetic trees. *Math. Biosci.* **53**: 131–147.

Rodrigo A. G., M. Kelly-Borges, P. R. Bergquist, and P. L. Bergquist. 1993. A randomization test of the null hypothesis that two cladograms are sample estimates of a parametric phylogenetic tree. *New Zealand J. of Botany* **31**: 257-268.

Rossello, F., and G. Valiente. 2006. The transposition distance for phylogenetic trees. *arXiv* q-bio/0604024v1 [q-bio.PE]

Ruta, M., and J. R. Bolt. 2008. The brachyopoid *Hadrokkosaurus bradyi* from the early Middle Triassic of Arizona, and a phylogenetic analysis of lower jaw characters in temnospondyl amphibians. *Acta Palaeontologica Polonica*, **53**: 579-592.

Ruta, M., and M. I. Coates. 2007. Dates, nodes and character conflict: Addressing the lissamphibian origin problem. *J. Syst. Palaeontology* **5**: 69–122.

Sánchez-Villagra, M. R., and B. A. Williams. 1998. Levels of homoplasy in the evolution of the mammalian skeleton. *J. Mamm. Evol.* **5**: 113–126.

Sánchez-Villagra, M. R., I. Horovitz, and M. Motokawa. 2006. A comprehensive morphological analysis of talpid moles (Mammalia) phylogenetic relationships. *Cladistics* **22**: 59–88.

Sanderson, M. J., and M. J. Donoghue. 1989. Patterns of variation in levels of homoplasy. *Evolution* **43**: 1781-1795.

Sansom, R. S., S. E. Gabbott, and M. A. Purnell. 2010. Non-random decay of chordate characters causes bias in fossil interpretation. *Nature* **463**: 797-800.

Sansom, R. S., S. E. Gabbott, and M. A. Purnell. 2011. Decay of vertebrate characters in hagfish and lamprey (Cyclostomata) and the implications for the vertebrate fossil record. *Proc. Roy. Soc. Lond. B Biol. Sci.* **1709**: 1150-1157.

Sansom, R. S., and M. A. Wills. 2013. Fossilization causes organisms to appear erroneously primitive by distorting evolutionary trees. *Scientific Reports* **3**: 2545.

Sansom, R. S. 2015. Bias and sensitivity in the placement of fossil taxa resulting from interpretations of missing data. *Syst, Biol.* **64**: 256-266.

Sereno, P. C., and S. L. Brusatte. 2008. Basal abelisaurid and carcharodontosaurid theropods from the Lower Cretaceous Elrhaz formation of Niger. *Acta Palaeontologica Polonica* **53**: 15–46.

Shimada, K. 2005. Phylogeny of lamniform sharks (Chondrichthyes: Elasmobranchii) and the contribution of dental characters to lamniform systematics. *Paleontol. Res.* **9**: 55–72.

1009 Sigurdsen, T., A. K. Huttenlocker, S. P. Modesto, T. B. Rowe, and R. Damiani. 2012. Reassessment of the
1010 morphology and paleobiology of the therocephalian Tetracynodon darti (Therapsida), and the
1011 phylogenetic relationships of Baurioidea. *Journal of Vertebrate Paleontology* **32**: 1113-1134.
1012 Simmons, N. B., and T. M., Conway. 2001. Phylogenetic relationships of mormoopid bats (Chiroptera:
1013 Mormoopidae) based on morphological data. *Bull. Am. Mus. Nat. Hist.* **258**: 1–100.
1014 Simmons, N. B., K. L. Seymour, J. Habersetzer, and G. F. Gunnell. 2008. Primitive early Eocene bat from
1015 Wyoming and the evolution of flight and echolocation. *Nature* **451**: 818–821.
1016 Skutschas, P. P., and Y. M. Gubin. 2012. A new salamander from the late Paleocene—Early Eocene of
1017 Ukraine. *Acta Palaeontologica Polonica* **57**: 135-148.
1018 Smith, N. D., and D. Pol. 2007. Anatomy of a basal sauropodomorph dinosaur from the early Jurassic
1019 Hanson formation of Antarctica. *Acta Palaeontologica Polonica* **52**: 657–674.
1020 Smith, N. D. 2010. Phylogenetic analysis of Pelecaniformes (Aves) based on osteological data: Implications
1021 for waterbird phylogeny and fossil calibration studies. *PLoS One* **5**: e13354.
1022 Smith, N. 2011. Taxonomic revision and phylogenetic analysis of the flightless Mancallinae (Aves, Pan-
1023 Alcidae). *ZooKeys* **91**: 1-116.
1024 Sokal, R. R., and P. H. A. Sneath. 1963. *Principles of Numerical Taxonomy*. Freeman and Company,
1025 London.
1026 Song, H., and S. R. Bucheli. 2010. Comparison of phylogenetic signal between male genitalia and non-
1027 genital characters in insect systematics. *Cladistics* **26**: 23–35.
1028 Sparks, J. S. 2008. Phylogeny of the cichlid subfamily Etroplinae and taxonomic revision of the Malagasy
1029 cichlid genus *Paretroplus* (Teleostei: Cichlidae). *Bull. Am. Mus. Nat. Hist.* **314**: 1**–**151.
1030 Spaulding, M., M. A. O'Leary, and J. Gatesy. 2009. Relationships of Cetacea (Artiodactyla) among
1031 mammals: Increased taxon sampling alters interpretations of key fossils and character evolution. *PLoS*
1032 *One* **4**: e7062+.
1033 Springer, M.S., A. Burk-Herrick, R. Meredith, E. Eizirik, E. Teeling, S. J. O'Brien, and W. J. Murphy. 2007.
1034 The adequacy of morphology for reconstructing the early history of placental mammals. *Syst. Biol.* **56**:
1035 673–684.
1036 Steel, M.A., and D. Penny. 1993. Distributions of tree comparison metrics – some new results. *Syst. Biol.* **42**:
1037 126–141.

1038 Sues, H.-D., and A. Averianov. 2009. A new basal hadrosauroid dinosaur from the Late Cretaceous of
1039 Uzbekistan and the early radiation of duck-billed dinosaurs. *Proc. Roy. Soc. Lond. B Biol. Sci.* **276**:
1040 2549–2555.
1041 Sues, H.-D., and R. R. Reisz. 2008. Anatomy and phylogenetic relationships of *Sclerosaurus armatus*
1042 (Amniota: Parareptilia) from the Buntsandstein (Triassic) of Europe. *J. Vertebr. Paleontol.* **28**: 1031–
1043 1042.
1044 Swartz, B. A. 2009. Devonian actinopterygian phylogeny and evolution based on a redescription of
1045 Stegotrachelus finlayi. *Zoo. J. Linn. Soc.* **156**: 750-784.
1046 Thompson, R. S, E. V. Barmann, and R. J. Asher. 2012. The interpretation of hidden support in combined
1047 data phylogenetics. *J. Zool. Syst. Evol. Res.* **50**: 251-263.
1048 Tsuji, L. A. 2013. Anatomy, cranial ontogeny and phylogenetic relationships of the pareiasaur Deltavjatia
1049 rossicus from the Late Permian of central russia. *Earth and Environmental Science Transactions of*
1050 *the Royal Society of Edinburgh* **104**: 81-122.
1051 Tsuji, L. A., and J. Mueller. 2009. Assembling the history of the Parareptilia: phylogeny,
1052 diversification, and a new definition of the clade. *Fossil Record* **12**: 71-81.
1053 Turner, A. H., and J. J. W. Sertich. 2010. Phylogenetic history of *Simosuchus clarki* (Crocodyliformes:
1054 Notosuchia) from the late cretaceous of Madagascar. *Journal of Vertebrate Paleontology* **30**: 177-236.
1055 Vallin, G., and M. Laurin. 2004. Cranial morphology and affinities of *Microbrachis*, and a reappraisal of the
1056 phylogeny and lifestyle of the first amphibians. *Journal of Vertebrate Paleontology* **24**: 56–72.
1057 Venczel, M. 2008. A new salamandrid amphibian from the Middle Miocene of Hungary and its phylogenetic
1058 relationships. *J. Syst. Palaeontology* **6**: 41–59.
1059 Vermeij, G. J., and S. J. Carlson. 2000. The muricid gastropod subfamily Rapaninae: Phylogeny and
1060 ecological history. *Paleobiology* **26**: 19–46.
1061 Vieira, G. H. C., G. R. Colli, and S. N. Bao. 2005. Phylogenetic relationships of corytophanid lizards (Iguania,
1062 Squamata, Reptilia) based on partitioned and total evidence analyses of sperm morphology, gross
1063 morphology, and DNA data. *Zoolog. Scripta* **34**: 605–625.
1064 Ward, C. V. 1997. Functional anatomy and phyletic implications of the hominoid trunk and hindlimb. Pp. 101-
1065 130 *in* B. Begun, C. Ward, and M. Rose (eds.). *Function, Phylogeny, and Fossils: Miocene Hominoid*
1066 *Evolution and Adaptations*. Plenum Press, New York.
1067 Waterman, M. S., and T. F. Smith. 1978. On the similarity of dendrograms. *J. Theor. Biol.* **73**: 789–800.

1068 Wheeler, W. 1999. Measuring topological congruence by extending character techniques. *Cladistics* **15**:
1069 131–135.
1070 Wiens, J. J. 2003a. Incomplete taxa, incomplete characters, and phylogenetic accuracy: Is there a missing
1071 data problem? *J. Vertebr. Paleontol.* **23**: 297–310.
1072 Wiens, J. J. 2003b. Missing data, incomplete taxa, and phylogenetic accuracy. *Syst. Biol.* **52**: 528–538.
1073 Wiens, J. J., and B. D. Hollingsworth. 2000. War of the iguanas: Conflicting molecular and morphological
1074 phylogenies and long-branch attraction in iguanid lizards. *Syst. Biol.* **49**: 143-159.
1075 Wilkinson, M. 1995. TAXEQ3: Safe taxonomic reduction using taxonomic equivalence. Computer program
1076 published by the author. http://www.softpedia.com/get/Science-CAD/TAXEQ3.shtml
1077 Williams, W. T., and H. T. Clifford. 1971. On the comparison of two classifications of the same set of
1078 elements. *Taxon* **20**: 519-522.
1079 Wills, M. A., R. A. Jenner, and C. Ni Dhubhghaill. 2009. Eumalacostracan evolution: conflict between three
1080 sources of data. *Arthropod Systematics and Phylogeny* **67**: 71-90.
1081 Wilson, M. V. H., and A. M. Murray. 2008. Osteoglossomorpha: phylogeny, biogeography, and fossil record
1082 and the significance of key African and Chinese fossil taxa. Pp. 185-219 *in* L. Cavin, A. Longbottom,
1083 and M. Richter (eds.). *Fishes and the Break-up of Pangaea*. Geological Society of London, Special
1084 Publications.
1085 Xu, G.-H., and K.-Q. Gao, 2011. A new scanilepiform from the Lower Triassic of northern Gansu Province,
1086 China, and phylogenetic relationships of non-teleostean Actinopterygii. *Zoo. J. Linn. Soc.* **161**: 595-
1087 612.
1088 Young, M. T., and M. B. de Andrade. 2009. What is *Geosaurus*?: Redescription of *Geosaurus giganteus*
1089 (Thalattosuchia: Metriorhynchidae) from the Upper Jurassic of Bayern, Germany. *Zool. J. Linn. Soc.*
1090 **157**: 551–585.
1091 Young, N. M. 2005. Estimating hominoid phylogeny from morphological data: Character choice, phylogenetic
1092 signal and postcranial data. Pp. 19-31 *in* D. E. Lieberman, R. J. Smith, and J. Kelley (eds.) *Interpreting*
1093 *the Past: Essays on Human, Primate and Mammal Evolution in Honor of David Pilbeam*. Brill
1094 Academic Publishers Inc., Boston.
1095 Zanno, L. E., D. D. Gillette, L. B. Albright, and A. L. Titus. 2009. A new North American therizinosaurid and
1096 the role of herbivory in 'predatory' dinosaur evolution. *Proc. Roy. Soc. Lond. B Biol. Sci.* **276**: 3505-
1097 3511.

Zaretskii, K. A. 1965. Constructing a tree on the basis of a set of distances between the hanging vertices. *Uspekhi Mat. Nauk* **20**: 90–92.

Table and figure legends

Author(s)	Year	Clade	Taxa	Cran. chars.	Post. chars.	Cran. miss %	Post. miss %	ILD	IRD NND	Cran. CI	Post. CI	Cran. ci	Post. ci	Cran. HER	Post. HER
Bourdon et al.	2009	Palaeognathae	17	35	94	4.0	8.0	1.000	0.979	0.956	0.855	0.971	0.908	0.975	0.921
Clarke & Middleton	2008	Avialae	20	45	141	38.7	22.7	0.001	0.127	0.740	0.643	0.774	0.732	0.734	0.703
de Pietri et al.	2011	Charadriiformes	12	15	25	0.6	2.3	0.158	0.685	0.667	0.758	0.567	0.756	0.383	0.442
Hospitaleche et al.	2007	Sphenisciformes	19	20	23	5.5	14.6	0.290	0.600	0.479	0.509	0.539	0.617	0.502	0.367
Manegold & Toepfer	2013	Picidae	27	20	47	1.7	0.4	0.485	0.005	0.710	0.458	0.816	0.647	0.904	0.761
Mauricio et al.	2012	Rhynocriptidae	38	40	46	1.1	1.1	0.043	0.048	0.480	0.517	0.646	0.654	0.721	0.710
Mayr et al.	2010	Basal Psittaciformes	32	25	68	3.9	3.3	0.466	0.180	0.456	0.341	0.541	0.456	0.400	0.355
Mayr	2010	Caprimulgiformes	10	18	51	1.7	2.4	0.572	0.374	0.692	0.643	0.735	0.735	0.604	0.522
Mayr	2011	Pelagornithidae	25	19	39	21.3	8.1	0.002	0.071	0.431	0.388	0.569	0.419	0.494	0.426
Nesbitt et al.	2011	Strisores	20	36	61	13.6	23.1	0.002	0.002	0.656	0.447	0.756	0.566	0.753	0.379
O'Connor et al.	2009	Enantiornithines	19	40	152	43.7	29.7	0.067	0.557	0.712	0.578	0.760	0.699	0.608	0.565
Smith	2010	Pelecaniformes	53	94	370	5.3	2.7	0.845	0.865	0.493	0.441	0.622	0.581	0.804	0.799
Smith	2011	Mancallinae	54	51	163	3.5	6.7	0.063	0.175	0.214	0.212	0.313	0.331	0.511	0.581
Allain & Aquesbi	2008	Sauropods	18	62	150	31.1	21.5	0.006	0.009	0.683	0.686	0.756	0.790	0.662	0.682
Allain et al.	2012	Spinosaurids	23	75	94	19.5	24.7	0.044	0.025	0.514	0.579	0.649	0.675	0.516	0.656
Andres et al.	2010	Basal Pterosaurs	18	44	31	18.2	17.7	0.054	0.556	0.667	0.574	0.718	0.649	0.658	0.537
Burns et al.	2011	Ankylosaurs	17	45	15	12.8	36.9	0.365	0.024	0.548	0.704	0.661	0.661	0.527	0.439
Butler et al.	2008	Ornithischia	29	124	88	19.8	25.0	0.367	0.134	0.502	0.521	0.666	0.693	0.574	0.604
Carrano & Sampson	2008	Ceratosauria	12	66	76	36.3	31.9	0.304	0.451	0.863	0.769	0.924	0.846	0.854	0.690
Ezcurra & Cuny	2007	Coelophysoidea	11	63	73	12.8	15.3	0.011	0.791	0.610	0.731	0.678	0.768	0.617	0.668
Gates & Sampson	2007	Hadrosaurinae	13	59	14	14.5	20.3	0.276	0.024	0.607	0.667	0.679	0.674	0.499	0.390
Godefroit et al.	2008	Hadrosauridae	15	30	7	9.3	8.5	1.000	0.003	0.918	1.000	0.958	0.956	0.927	1.000
Lu et al.	2010	Pterosauria	37	59	55	8.5	21.0	0.002	0.066	0.561	0.457	0.695	0.572	0.698	0.624
Lu et al.	2013	Oviraptoridae	17	108	69	31.1	33.2	0.002	0.001	0.756	0.516	0.837	0.581	0.760	0.315
Martinez & Alcober	2009	Dinosauria	12	28	70	33.6	27.7	0.067	0.779	0.732	0.560	0.780	0.677	0.544	0.251
Nesbitt et al.	2009	Early Dinosaurs	35	113	194	32.2	30.8	0.001	0.003	0.413	0.469	0.540	0.603	0.429	0.653
Nesbitt	2011	Archosaurs	55	166	223	23.1	28.2	0.001	0.030	0.402	0.413	0.576	0.562	0.592	0.706
Osi & Makadi	2009	Ankylosauria	15	43	14	20.3	32.9	0.117	0.216	0.537	0.643	0.626	0.598	0.502	0.827
Sereno	2008	Carcharodontosaurids	9	37	23	31.8	29.5	0.485	0.327	0.848	0.639	0.899	0.692	0.950	0.201
Smith & Pol	2007	Plateosauria	26	115	232	23.2	17.9	0.250	0.414	0.461	0.362	0.550	0.533	0.563	0.581
Sues & Averianov	2009	Iguanodontidae	25	100	38	11.9	24.4	0.265	0.014	0.650	0.672	0.280	0.305	0.838	0.827
Zanno et al.	2009	Therizinosauria	21	80	191	29.3	25.1	0.054	0.380	0.528	0.524	0.617	0.609	0.460	0.540
Asher et al.	2005	Lagomorpha	29	157	58	7.1	1.8	0.001	0.048	0.373	0.331	0.460	0.403	0.501	0.429
Asher & Hofreiter	2006	Afrotheria	20	90	22	9.0	4.4	0.795	0.617	0.406	0.441	0.487	0.445	0.404	0.325
Asher	2007	Eutheria	46	121	75	6.4	5.2	0.001	0.040	0.223	0.275	0.305	0.406	0.286	0.384
Beard et al.	2009	Amphipithecidae	22	216	55	17.0	37.6	0.004	0.478	0.305	0.569	0.503	0.555	0.426	0.427
Beck et al.	2008	Marsupiala	27	101	139	10.7	9.1	0.004	0.259	0.432	0.400	0.532	0.517	0.543	0.477
Bloch et al.	2007	Plesiadapiforms	12	98	61	3.2	17.3	0.001	0.062	0.401	0.627	0.531	0.702	0.337	0.295
Gaubert et al.	2005	Feliformia	39	229	100	1.1	3.1	0.001	0.500	0.356	0.335	0.266	0.219	0.497	0.519
Gaudin et al.	2009	Pholidota	16	298	90	14.8	23.3	0.528	0.155	0.339	0.325	0.598	0.658	0.522	0.532
Lister et al.	2005	Cervidae	10	21	41	0.5	1.7	0.729	0.410	0.750	0.623	0.778	0.659	0.418	0.261
Martinelli & Rougier	2007	Ictidosauria	16	65	17	10.9	31.3	0.088	0.001	0.674	0.770	0.790	0.778	0.646	0.837
Phillips et al.	2009	Monotremes	55	325	126	25.3	21.0	0.004	0.002	0.337	0.503	0.543	0.625	0.682	0.811
Pine et al.	2012	Oryzomyini	36	55	34	3.1	16.5	0.033	0.207	0.271	0.384	0.371	0.388	0.248	0.304
Pujos et al.	2007	Folivora	18	25	17	7.8	21.6	0.066	0.441	0.455	0.628	0.503	0.610	0.382	0.457
Sanchez-Villagra et al.	2006	Talpidae	17	73	83	4.7	4.3	0.003	0.010	0.495	0.549	0.587	0.652	0.445	0.609
Sigurdsen et al.	2012	Baurioidea	20	95	20	8.3	35.8	0.507	0.049	0.520	0.706	0.638	0.725	0.588	0.597
Simmons & Conway	2001	Mormoopidae	16	103	106	16.0	14.0	1.000	0.974	0.650	0.659	0.539	0.586	0.679	0.685
Simmons et al.	2008	Chiroptera	25	48	154	16.9	21.7	0.201	0.559	0.445	0.419	0.524	0.521	0.525	0.459
Spaulding et al.	2009	Artiodactyla	37	353	137	17.6	21.0	0.018	0.711	0.235	0.307	0.326	0.444	0.288	0.444
Brochu et al.	2010	Crocodyloidea	24	57	23	7.3	14.1	0.032	0.030	0.598	0.540	0.703	0.530	0.658	0.464
Cheng et al.	2012	Sauropterygia	35	69	70	10.7	14.0	0.463	0.830	0.401	0.405	0.503	0.471	0.471	0.586
Gaffney et al.	2009	Bothremydidae	23	100	50	18.0	27.1	0.001	0.186	0.686	0.727	0.789	0.800	0.733	0.729
Hill	2005	Amniota	19	124	150	1.4	4.6	0.457	0.616	0.550	0.620	0.655	0.729	0.664	0.737
Ji et al.	2013	Ichthyosauria	23	42	56	27.4	16.9	0.024	0.001	0.620	0.716	0.655	0.808	0.653	0.822
Lee & Scanlon	2002	Serpentes	21	204	53	9.9	4.3	0.882	0.209	0.505	0.522	0.627	0.569	0.548	0.472
Li et al.	2007	Squamata	23	212	97	8.1	15.1	0.099	0.389	0.536	0.511	0.644	0.614	0.591	0.447
Lyson & Joyce	2009	Baenidae	10	31	17	11.2	14.7	0.960	0.155	0.672	0.708	0.753	0.755	0.503	0.571
Matsumoto et al.	2009	Choristodera	14	49	29	14.6	13.3	0.002	0.728	0.647	0.571	0.704	0.558	0.611	0.424
Muller & Reisz	2006	Eureptiles	24	58	19	12.3	16.3	0.171	0.028	0.418	0.468	0.479	0.309	0.347	0.399
Palci et al.	2013	Serpentes	26	181	29	15.4	13.4	0.613	0.945	0.468	0.595	0.594	0.639	0.570	0.662
Parilla-Bel et al.	2013	Metriorhynchoidea	29	139	59	25.0	38.3	0.191	0.005	0.582	0.862	0.692	0.883	0.743	0.918
Sues & Reisz	2008	Parareptilia	12	34	9	10.3	38.8	0.034	0.004	0.731	0.750	0.797	0.657	0.738	0.631
Tsuji & Mueller	2013	Pareiasauria	21	64	50	22.3	36.3	0.118	0.066	0.755	0.793	0.811	0.847	0.772	0.795
Vieira et al.	2005	Corytophanidae	12	40	28	0.8	0.3	0.485	0.625	0.788	0.673	0.829	0.745	0.819	0.597
Anderson et al.	2008	Batrachia	52	143	76	18.4	29.3	0.001	0.911	0.254	0.285	0.347	0.367	0.407	0.379
Holland & Long	2009	Tetrapodomorpha	12	54	41	8.5	21.3	0.009	0.093	0.800	0.710	0.843	0.774	0.777	0.655
Ruta & Coates	2007	Lissamphibia	46	189	104	20.9	15.7	0.001	0.196	0.320	0.314	0.466	0.470	0.460	0.463
Skutchas & Gubin	2012	Salamanders	21	16	40	18.5	34.2	0.067	0.015	0.565	0.491	0.512	0.558	0.283	0.276
Vallin & Laurin	2004	Stegocephalia	47	99	60	15.8	26.7	0.002	0.028	0.355	0.417	0.435	0.538	0.584	0.544
Venczel	2008	Caudata	15	19	16	1.8	12.5	1.000	0.145	0.538	0.783	0.596	0.854	0.516	0.876
Choo	2011	Actinopterygii	27	51	26	12.4	16.4	0.635	0.460	0.424	0.448	0.550	0.600	0.464	0.451
Dutel et al.	2012	Actinistia	30	83	23	29.1	22.1	0.904	0.288	0.424	0.500	0.566	0.548	0.523	0.442
Friedman	2007	Actinistia	25	136	40	25.0	36.6	0.003	0.150	0.464	0.523	0.631	0.691	0.580	0.511
Friedman	2008	Pleuronectiformes	19	16	42	19.4	8.9	0.284	0.223	0.548	0.526	0.599	0.595	0.423	0.608
Hilton & Forey	2009	Acipenseriformes	18	31	17	6.5	6.5	0.661	0.010	0.634	0.724	0.751	0.740	0.670	0.739
Hurley et al.	2007	Actinopterygii	24	53	16	19.3	17.4	0.484	0.001	0.506	0.562	0.632	0.606	0.562	0.421
Inamura et al.	2005	Cottoidei	8	16	15	0.0	0.0	0.353	0.273	0.640	0.700	0.656	0.810	0.390	0.601
Lopez-Arbarello	2011	Ginglymodi	29	65	19	24.0	15.4	0.321	0.071	0.454	0.471	0.612	0.489	0.576	0.477
Parenti	2008	Adrianichthyidae	31	37	43	0.6	8.8	0.574	0.565	0.677	0.584	0.844	0.644	0.778	0.688
Shimada	2005	Lamniformes	16	40	15	12.7	1.3	0.214	0.284	0.525	0.810	0.552	0.769	0.551	0.860
Sparks	2008	Etroplinae	25	25	55	0.8	4.0	0.593	0.474	0.867	0.707	0.890	0.833	0.962	0.925
Swartz	2009	Actinopterygii	22	28	31	20.5	21.7	0.309	0.088	0.552	0.581	0.705	0.727	0.472	0.607
Wilson & Murray	2008	Osteoglossomorpha	24	50	37	17.6	23.2	0.769	0.820	0.522	0.571	0.614	0.603	0.593	0.591
Xu & Gao	2011	Scanilepiformes	22	45	27	21.3	16.5	0.210	0.110	0.509	0.604	0.680	0.716	0.470	0.587

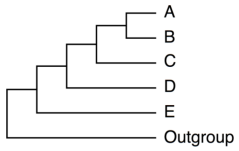
Table 1. Summary statistics for the 85 vertebrate morphological data sets analysed herein. Data set dimensions (numbers of taxa (ntax) and informative characters (nchar)) refer to the pre-processed matrices after applying safe taxonomic deletion rules (see text for details). 'Cran. char.' and 'Post. char' denote characters per partition. 'Cran. miss %' and 'Post. miss %' report the percentage of missing data cells for partitions. 'ILD' column reports the p-value resulting from an incongruence length difference test with 999 random partitions. IRD columns report the p-values resulting from incongruence relationship difference tests with 999 random partitions. 'IRD_{NND}' denotes the results of the IRD test using the Robinson Foulds tree-to-tree distance for nearest neighbouring trees. CI columns give ensemble consistency indices for partitions of the data set (craniodental or postcranial). Similarly, ci columns report the mean per character consistency indices for partitions of the data set (craniodental or postcranial) when optimised onto the globally most parsimonious trees. HER gives the homoplasy excess ratio for partitions of the data set (craniodental or postcranial) derived from 999 randomized matrices. An expanded version of this table containing additional statistics is provided within the Supporting Information.

Taxonomic Group	Aves	Other Ornithodira	Synapsida	Other Reptiles	Amphibians	Fishes	Log likelihood ratio (G)	*p*
No. data sets	13	19	18	15	6	14		
Consistency indices sig. diff.	8	8	9	7	4	4	4.283	0.51
ILD	4	7	10	5	4	1	11.808	0.038
IRD NND	3	9	6	5	2	2	4.801	0.441
IRD MR	1	7	5	3	0	1	9.522	0.09

Table 2. Number of data matrices in higher taxonomic groups, with a tally of those with significant ($p < 0.05$) results for various partition homogeneity tests. G-test and *p* values quoted for the null that significant results are equally likely in the five higher taxonomic groups in each case.

a

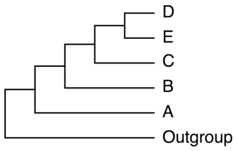
Taxon	Characters		
	1	2	3
Outg.	0	0	0
A	1	1	1
B	1	1	1
C	0	1	1
D	0	0	1
E	0	0	0



Char.	CI	RI	RC
1	1.000	1.000	1.000
2	1.000	1.000	1.000
3	1.000	1.000	1.000

b

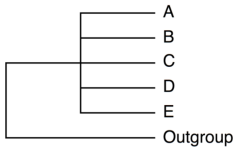
Taxon	Characters		
	4	5	6
Outg.	0	0	0
A	0	0	0
B	0	0	1
C	0	1	1
D	1	1	1
E	1	1	1



Char.	CI	RI	RC
4	1.000	1.000	1.000
5	1.000	1.000	1.000
6	1.000	1.000	1.000

c

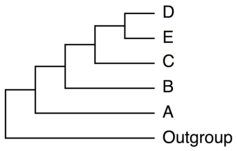
Taxon	Characters					
	1	2	3	4	5	6
Outg.	0	0	0	0	0	0
A	1	1	1	0	0	0
B	1	1	1	0	0	1
C	0	1	1	0	1	1
D	0	0	1	1	1	1
E	0	0	0	1	1	1



Char.	CI	RI	RC
1	0.875	0.750	0.750
2	0.750	0.750	0.625
3	0.625	0.250	0.250
4	0.875	0.750	0.750
5	0.750	0.750	0.625
6	0.625	0.250	0.250

d

Taxon	Characters							
	1	2	3	4	5	6	5'	6'
Outg.	0	0	0	0	0	0	0	0
A	1	1	1	0	0	0	0	0
B	1	1	1	0	0	1	0	1
C	0	1	1	0	1	1	1	1
D	0	0	1	1	1	1	1	1
E	0	0	0	1	1	1	1	1



Char.	CI	RI	RC
1	0.500	0.000	0.000
2	0.500	0.500	0.250
3	0.500	0.000	0.000
4	1.000	1.000	1.000
5	1.000	1.000	1.000
6	1.000	1.000	1.000
5'	1.000	1.000	1.000
6'	1.000	1.000	1.000

Figure 1. (a and b) Characters sampled from different anatomical regions can yield radically different most parsimonious trees (MPT) when analysed in isolation. In both cases, there is no homoplasy *within* either region (characters 1-3 or characters 4-6), and a single MPT results in each case. (c) Combining the data from both partitions (characters 1-6) yields four MPTs, the strict consensus of which (illustrated) is entirely unresolved. Character statistics have been averaged over the four trees. (d) Two additional characters (5' and 6') are sampled from the same region as 'b', and these have the same distribution as 5 and 6 respectively. Analysis of all characters now reveals a single MPT with relationships identical to those in 'b' (characters 4-6). Characters 4-6 , 5' and 6' contain no homoplasy: all conflicts are resolved with a cost to characters 1-3. In this case, the MPT is identical to the result that would be obtained by a clique analysis (*sensu* Le Quesne 1969).

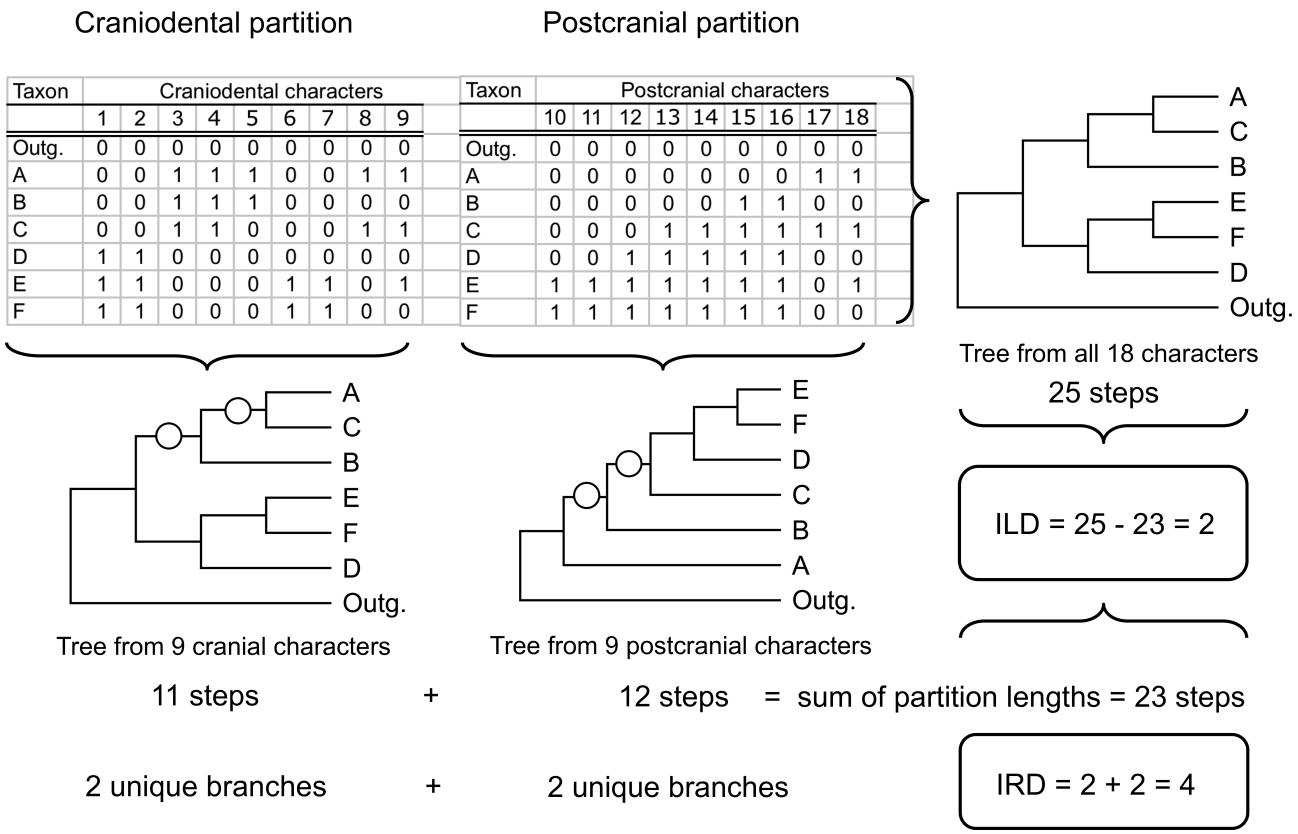


Figure 2. Calculation of two partition inhomogeneity metrics for 'craniodental' and 'postcranial' partitions of a hypothetical data set. In this example there are equal numbers of craniodental (1-9) and postcranial (10-18) characters, but this need not be the case. For the Incongruence Length Difference (ILD) measure, maximally parsimonious trees (MPTs) are inferred from the craniodental and postcranial partitions of the data *independently*. The summed lengths of these trees (11 steps + 12 steps) is the sum of partition lengths (23 steps). In parallel with this, an MPT is inferred from both partitions analysed *simultaneously*. This tree is longer (25 steps) than the sum of partition lengths (23 steps), and the difference between them is the ILD (25 – 23 = 2). The ILD represents the reduction in homoplasy afforded by the isolation of the two partitions (two extra steps are needed when the partitions are combined). For the Incongruence Relationship Difference (IRD) measure, the branching structure of the craniodental and postcranial partition MPTs are compared (rather than their lengths) using one of several possible tree-to-tree distance metrics. Here, we illustrate the symmetric difference distance (RF) of Robinson and Foulds (1991). Open circles mark branches in either the craniodental or postcranial MPT that are absent from the other. The tally of these unique branches on both trees is the RF (2 + 2 = 4). Some background level of ILD or IRD is anticipated wherever a data set contains homoplasy. In order to interpret these observed metrics, therefore, we need to know what values would be expected for partitions of similar data sets in similar proportions. Random character partitions are used to

generate null distributions for both the ILD and IRD, and observed values deemed significantly different from the null if they lie in some specified fraction of the tails.

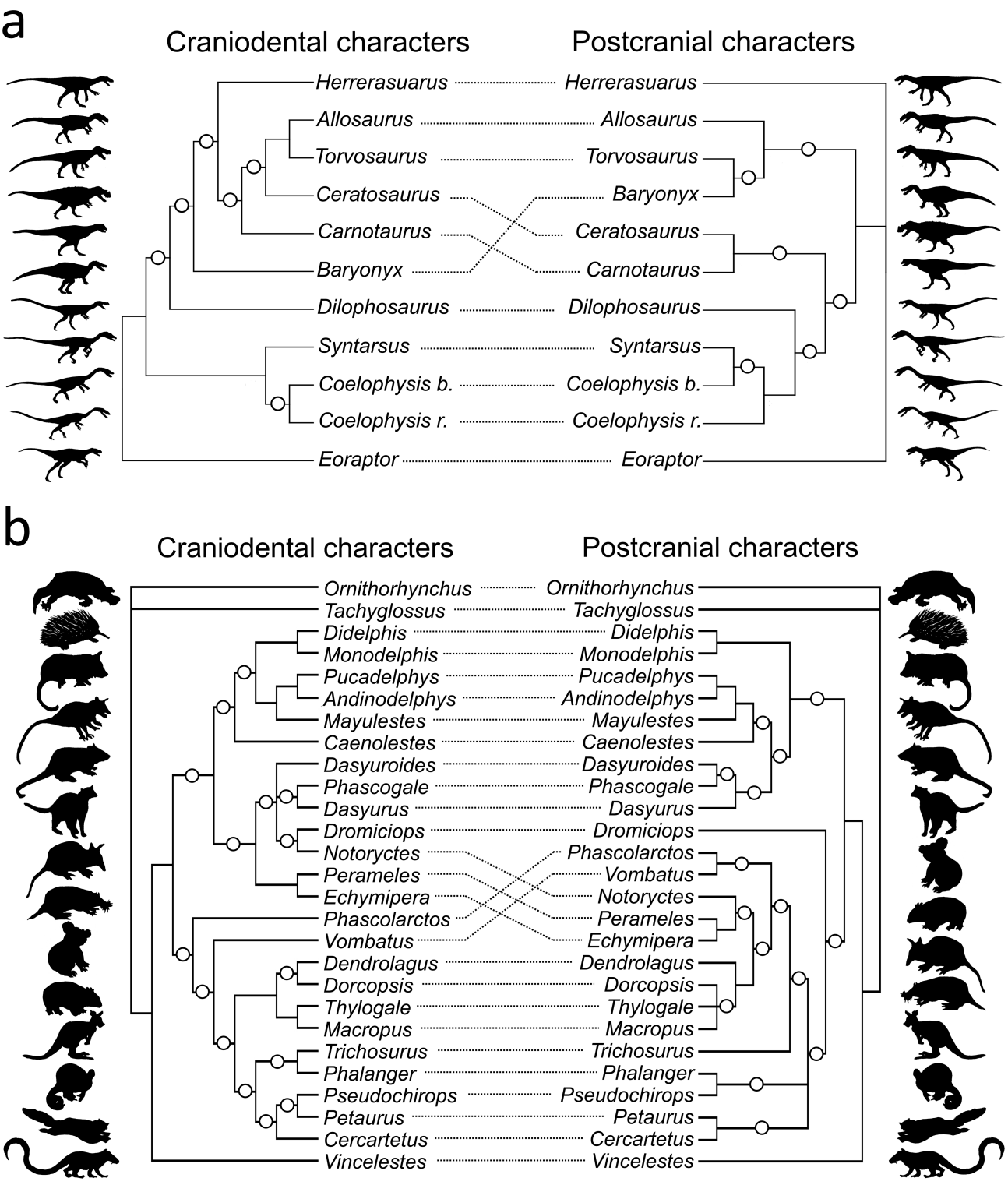


Figure 3. Most parsimonious trees derived from craniodental and postcranial partitions can be significantly more different than we would expect. Tanglegrams computed using *Dendroscope* (Huson and Scornavacca,

2012). (a) The theropod data of Ezcurra and Cuny (2007) yielded one most parsimonious tree from the craniodental partition and six from the postcranial partition, the latter summarized as a majority rule tree merely for ease of visualization (we advocate the use of tests based upon mean nearest neighbours within sets of most parsimonious trees). Branches labelled with circles are unique to one or other tree (those unlabelled are common to both). The Robinson Foulds (RF) distance between the two is simply the sum of unique branches (6+6=12). While the ILD test returned a highly significant result (p = 0.011) and our new incongruence relationship difference test (IRD_{NND}) using RF did not (p = 0.791). (b) Six craniodental and four postcranial trees in the mammalian data of Beck (2008), again summarized as majority rule consensus trees for visualization. In this case, the incongruence length difference (ILD) test for partition homogeneity returned a highly significant result (p = 0.004) whereas our incongruence relationship difference (IRD_{NND}) test did not (p = 0.259). Indicative images are, from top to bottom on right hand side: *Ornithorhynchus*, *Tachyglossus*, *Didelphis*, *Monodelphis*, *Caenolestes*, *Dasyurus*, *Phascolarctos*, *Vombatus*, *Perameles*, *Echymipera*, *Macropus*, *Phalanger*, *Petaurus*, *Vincelestes*. Image of *Tachyglossus* courtesy of echidnasclub.com. See text for further explanation.

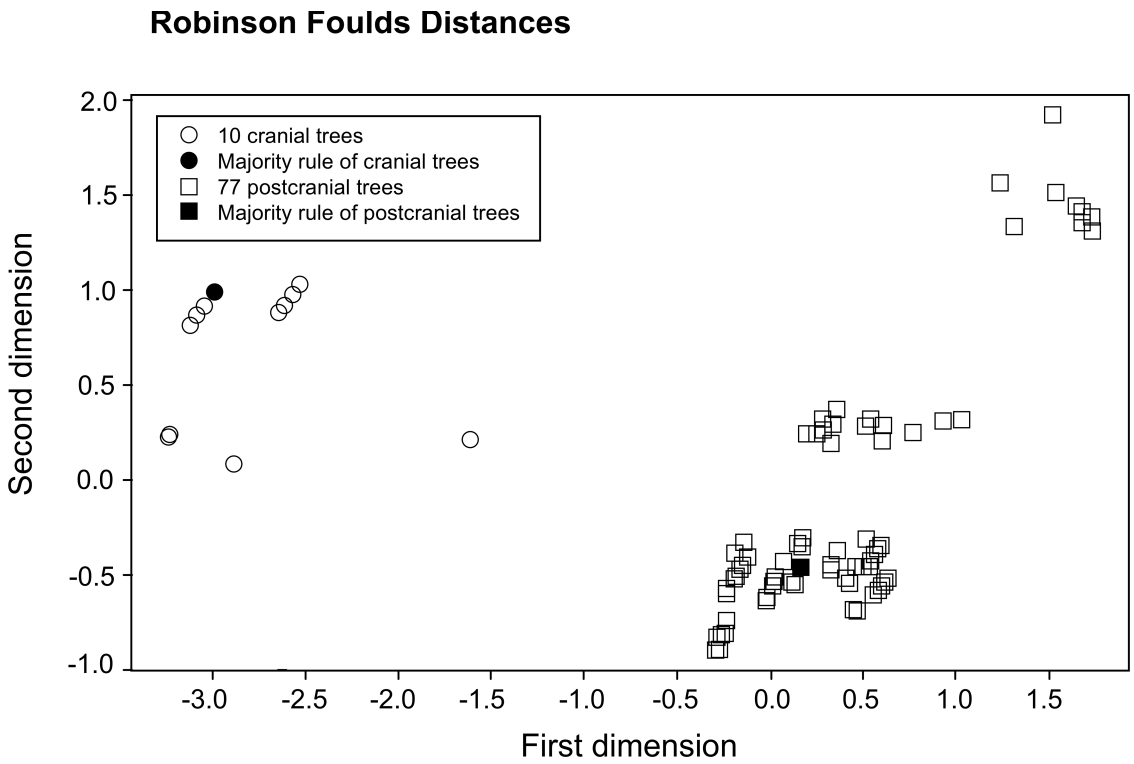


Figure 4. Summarising sets of most parsimonious trees (MPTs) for partitions prior to calculating tree-to-tree distances is computationally much faster than calculating distances between nearest neighbours. However, majority rule trees present modal or most frequent relationships, and may therefore plot eccentrically in tree

space. This is an undesirable property when attempting to summarise distances between sets of trees. Figure shows tree-to-tree distances for craniodental and postcranial partitions of the mammalian data of Pujos (2007). Distance matrices have been plotted in two dimensions using non-metric multidimensional scaling (NMDS), and rotated using principal components analysis (PCA). Circles indicate craniodental trees and squares indicate postcranial trees. Open symbols denote original MPTs, filled symbols (black) denote majority rule trees.

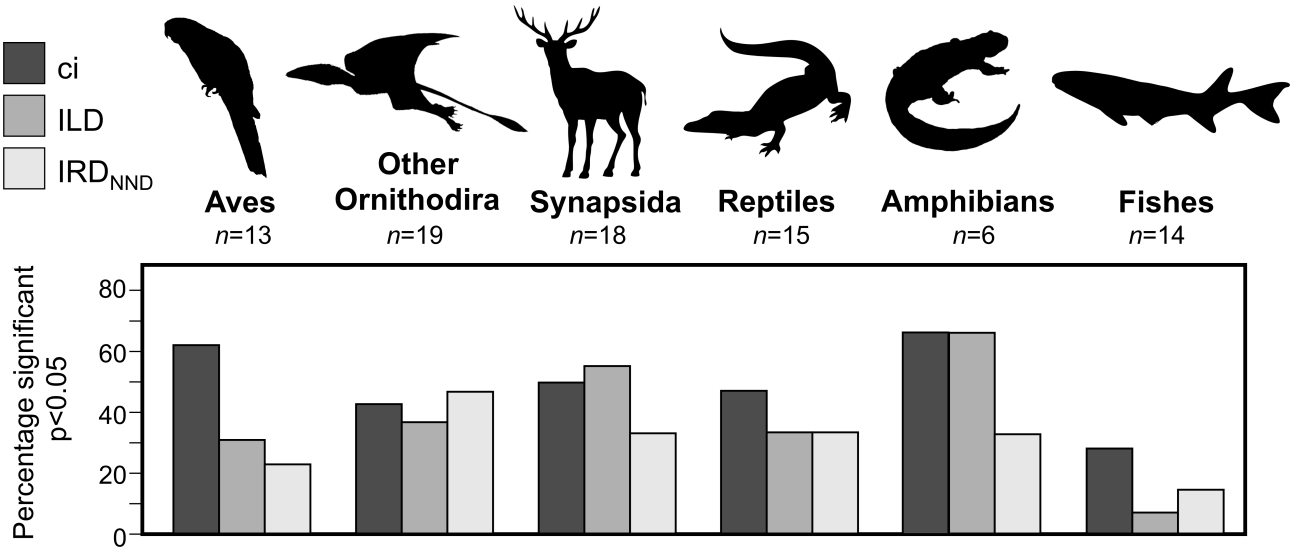


Figure 5. Summary of results from three partition homogeneity tests, subdivided by higher taxonomic group. Black bars denote significant ($p < 0.05$) differences in ci values between partitions (Mann-Whitney U tests). Gray bars denote significant ILD results, while light gray bars indicate significant IRD_{NND} results. Images from phylopic.org (courtey of Michael Scroggie, Oscar Sanisidro, B. Kimmel and Steven Coombs). Salamander and pterosaur courtesy of Matt Reinbold (modified by T. Michael Keesey) and Mark Witton. (http://creativecommons.org/licenses/by-sa/3.0/).